PROGRAMME DE RECHERCHE SUR LES NEURONES OPTIQUES APPLIQUÉS À LA RECONNAISSANCE DE CIBLES

INO #94-7404-2RC RET N/A

INO NOI

RAPPORT FINAL

PROGRAMME DE RECHERCHE SUR LES NEURONES OPTIQUES APPLIQUÉS À LA RECONNAISSANCE DE CIBLES

INO #94-7404-2RC RET N/A

Présenté à :

Centre de Recherche pour la Défense, Valcartier
2459 boul. Pie XI Nord
Courcelette, Québec G0A 1R0
Autorité scientifique : Paul Chevrette

Préparé par :

Henri H. Arsenault, Ph.D.
Alain Bergeron, Ph.D.
Nichola Desnoyers, M.Sc.
Erik Eustache, M.Sc.
Denis Gingras, Ph.D.
Donald Prévost, Ph.D.
Yunlong Sheng, Ph.D.
Michel Arsenault, M.Sc.
Sylvain Deschênes, M.Sc.
Michel Doucet, M.Sc.
Jonny Gauvin, M.Sc.
Sébastien Parent, M.Sc.
Roland Schaefer, Ph.D.
Luc Veilleux, Tech.

Préparé par :

Michel Doucet
Responsable de projet

Autorisé par :

Denis Gingras
Directeur
Systèmes numériques et optiques

INSTITUT NATIONAL D'OPTIQUE
369, rue Franquet
Sainte-Foy (Québec) G1P 4N8

Septembre 1997

Defence Research and Development Canada | Recherche et développement pour la défense Canada

Canada

TITLE:(U) PROGRAMME DE RECHERCHE SUR LES NEURONES OPTIQUES APPLIQUES A LA RECONNAISSANCE DE CIBLES

PERSONAL AUTHOR: Doucet, M.

CORPORATE AUTHOR: Defence Research Establishment Valcartier, Valcartier QUE (CAN);Institut National d'Optique, Ste-Foy QUE (CAN)

ABSTRACT: (U) Le projet decrit dans ce rapport porte sur la reconnaissance automatique en temps reel d'objets, et plus specifiquement de cibles militaires (RAC), a partir de leurs images. L'objectif du projet consiste a illustrer a l'aide d'un modele d'essai, les capacites du traitement optique de l'information et des reseaux de neurones dans le domaine de la RAC.

DESCRIPTORS: (U) Reconnaissance;Fiber optics;Detectors;Real time operations;Experimental data;Artificial intelligence;Nerve cells;Neural nets;Pattern recognition;Algorithms

REPORT NUMBER: DREV-CR-973;INO-94-7404-2RC; Final Report

REPORT DATE: 01 Sep 1997

NUMBER OF PAGES : 267

DRDKIM N°: 98-02206

CANDIS N°: 509058

Format(s): Hardcopy;Document Image stored on Optical Disk

INO NOI

Table de matières

Sommaire ..... **vii**

CHAPITRE 1 ..... **11**

1.1 Introduction **11**

1.2 Problématique et approches **13**

1.2.1 Énoncé du problème de RAC 13

1.2.2 Éléments de solution 15

1.2.3 Description des méthodes étudiées 16

1.2.3.1 Coefficients d'harmoniques circulaires (CHC) 17

1.2.3.2 Sections en pointe [2] 20

1.2.3.3 Descripteurs de Fourier [3] 21

1.2.3.4 Méthode des DCCF [4] 22

1.3 Sommaire de la première partie du projet **24**

1.3.1 Simulations numériques 24

1.3.2 Montages optiques 25

1.3.3 Traitement numérique 27

1.4 Sommaire de la seconde partie du projet **28**

1.4.1 Étude sur les méthodes des sections en pointe et des descripteurs de Fourier 28

1.4.1.1 Méthode des sections en pointe 28

1.4.1.2 Méthode des descripteurs de Fourier 30

1.4.2 Travaux sur les CHC 31

1.4.2.1 Nouvelle définition des CHC 31

1.4.2.2 Problème considéré pour le modèle d'essai 32

1.4.2.3 Systèmes optiques pour l'extraction des CHC 34

1.4.2.4 CHC angulaires 38

1.4.2.5 Résultats avec les CHC et les CHCA 39

1.4.2.6 Modèle d'essai 40

1.5 Conclusion **41**

PREMIERE PARTIE : CONTRIBUTION DE L'INO

CHAPITRE 2.....43

2.1 Signatures CHC..... 43

2.1.1 Nouvelle définition des CHC.....43

2.1.2 Formulation pour la mise en oeuvre optique.....45

2.1.3 Effet de la rotation de l'objet sur les valeurs C'_{nm} et S'_{nm}.....46

2.1.4 Introduction des CHC angulaires.....48

2.2 Incertitudes liées à l'évaluation optique des CHC..... 49

2.2.1 Évaluation de la borne supérieure de l'incertitude relative.....50

2.2.2 Ajustement de l'orientation des rosettes complémentaires.....53

2.3 Généralisation de la formulation des CHC optiques..... 55

2.3.1 Reformulation du calcul des CHC.....55

2.3.2 Caractérisation du système optique.....56

CHAPITRE 3.....59

3.1 Système à réseau de diffraction..... 59

3.1.1 Principe de fonctionnement.....60

3.1.2 Construction du montage à réseau de diffraction.....63

3.1.2.1 Principaux composants du montage.....63

3.1.3 Assemblage du montage.....71

3.1.3.1 Mécanique.....72

3.1.3.2 Alignement.....74

3.2 Montage à deux ACL..... 76

3.3 Appréciation globale des deux montages optiques..... 79

3.3.1 Système à réseau de diffraction.....79

3.3.2 Montage à deux ACL.....81

CHAPITRE 4.....83

4.1 Préparation des images..... 83

4.1.1 La scène.....85

4.1.2 L'acquisition.....86

4.1.3 Filtrage.....87

4.1.4 Segmentation.....88

4.1.5 Recentrage.....89

4.1.6 Affichage 91
4.1.7 Environnement de travail 92
4.1.8 Résultats de la préparation des images 92
4.2 Post-traitement 95
4.2.1 Acquisition des données 95
4.2.2 Corrections des données 98
4.2.3 Calcul des signatures CHC 98
4.2.4 Classification par réseaux de neurones 99

CHAPITRE 5 101
5.1 Résultats numériques 101
5.1.1 Classification 2D 102
5.1.1.1 Signatures CHC et CHC angulaires 102
5.1.1.2 Reconnaissance et discrimination 111
5.1.2 Classification 3D 113
5.1.2.1 Signatures CHC et CHC angulaires 113
5.1.2.2 Discrimination et reconnaissance 114
5.2 Résultats optiques 116
5.2.1 Reconnaissance et discrimination 2D 117
5.2.1.1 CHC angulaires 119
5.2.1.2 CHC 125
5.3 Commentaires généraux sur les CHC et CHCA 128
Références 131
ANNEXE A 133
A.1 Mécanique 133
A.2 Alignement 140

SECONDE PARTIE : CONTRIBUTION DU COPL

PREMIÈRE MÉTHODE :

6. INTRODUCTION 149

6.1. TÂCHES À ACCOMPLIR 150
6.2. PERSONNEL 150
6.3. REVUE DES PRINCIPALES TECHNIQUES 151
6.4. TECHNIQUES NON RETENUES 151
6.5. TECHNIQUES RETENUES 152
6.6. DESCRIPTION DE LA BANQUE D'IMAGES UTILISÉE 153

7. DESCRIPTION DE LA MÉTHODE DES TRAJECTOIRES D'ESPACE CARACTÉRISTIQUE AVEC RÉSEAU DE NEURONES (FST NN) 157

7.1. SEGMENTATION 157
7.2. SPECTRE DE PUISSANCE 158
7.3. COMPRESSION D'IMAGE 158
7.4. RÉSEAU DE NEURONES 158
7.5. TRAJECTOIRE D'ESPACE CARACTÉRISTIQUE 158
7.6. RECONNAISSANCE 159

8. RÉSULTATS DE LA MÉTHODE FST NN 161

8.1. INVARIANCE SOUS ÉCHELLE 162
8.1.1 Normalisation des vecteurs 162
8.1.2 Filtrage moyen 164
8.1.3 La qualité de l'invariance sous échelle 165
8.2. ROBUSTESSE FACE AU BRUIT 179
8.2.1 Bruit Blanc 180
8.2.2 Bruit de fond 181
8.3. CAPACITÉ DE RECONNAISSANCE 187
8.3.1 Réseau de neurones 187
8.3.2 Amélioration de l'apprentissage 188

SECONDE MÉTHODE :

9. INTRODUCTION 197

10. REHAUSSEMENT DES IMAGES 201

10.1. IMAGE ORIGINALE 201
10.2. ÉGALISATION DE L'HISTOGRAMME 202
10.3. REHAUSSEMENT DU CONTRASTE 203
10.4. FILTRAGE MÉDIAN 203

11. EXTRACTION DU CONTOUR AVEC LA TRANSFORMÉE EN ONDELETTES 207

11.1. MÉTHODES CONVENTIONNELLES 207
11.1.1 Filtre de Prewitt, Sobel et Frei-Chen 207
11.1.2 Détecteur de bords optimal 208
11.2. TRANSFORMÉE EN ONDELETTE 209
11.2.1 Réalisation de la transformée en ondelette numérique 210
11.2.1.1 Transformée en ondelette unidimensionnelle 210
11.2.1.2 Transformée en ondelettes d'une image bidimensionnelle 212
11.2.1.3 Ondelette de Daubechies à quatre coefficients 214
11.2.1.4 Ondelette de Daubechies à douze coefficients. 215
11.2.1.5 Ondelette de spline quadratique. 215
11.2.2 Avantages de la détection de bords avec la transformée en ondelette 218
11.2.3 Maxima locaux dans la direction du gradient. 219
11.3. RECHERCHE DU CONTOUR DE L'OBJET À ANALYSER. 220
11.3.1 Segmentation du plateau 220
11.3.2 Seuillage 220
11.3.3 Tracé de courbes 221
11.3.4 Problèmes rencontrés 222

12. DESCRIPTION DU CONTOURS 225

12.1. CODAGE EN CHAÎNE 225
12.2. DESCRIPTEUR DE FOURIER-MELLIN 226
12.3. DESCRIPTEUR DE FOURIER (DF) 227
12.3.1 Evaluation des descripteurs de Fourier 228
12.3.1.1 Elimination des points isolés et redondants 228
12.3.1.2 Calcul des descripteurs de Fourier 228
12.3.1.3 Normalisation des descripteurs de Fourier 230

13. RECONNAISSANCE D'OBJET 3-D 233

13.1. MÉTHODE DE DESSIN DE LIGNES 233
13.2. DESCRIPTEUR DE FOURIER INVARIANT SOUS TRANSFORMATION AFFINE 234
13.3. RÉSEAUX DE NEURONES 234
13.3.1 Modèle de transformation et matching 235
13.3.2 Réseau généralisé de fonctions de base 235
13.3.3 Modèle de trajectoire dans l'espace des caractéristiques (TEC) 237
13.3.3.1 Description de la méthode 237
13.3.3.2 Résultats obtenus 238

14. CONCLUSION 247

15. CONCLUSION GÉNÉRALE (COPL) 251

CONCLUSION GÉNÉRALE 257

Sommaire

Le projet décrit dans ce rapport porte sur la reconnaissance automatique en temps réel d'objets, et plus spécifiquement de cibles militaires (RAC), à partir de leurs images. L'objectif du projet consiste à illustrer à l'aide d'un modèle d'essai, les capacités du traitement optique de l'information et des réseaux de neurones dans le domaine de la RAC.

Le projet d'une durée de plus de trois ans a permis de mener des études sur différentes méthodes de RAC. Les efforts ont porté sur quatre méthodes dont trois appartiennent à la famille des approches dites à extraction de caractéristiques, tandis que la quatrième fait partie des méthodes basées sur la corrélation. Les approches à extraction de caractéristiques visent à faire la reconnaissance des objets par le biais d'un ensemble restreint de leurs caractéristiques morphologiques. On entend par caractéristique morphologique, une valeur numérique caractérisant l'image dans sa structure spatiale. La reconnaissance se fait sur l'ensemble ordonné des caractéristiques, lequel constitue une représentation mathématique, ou signature, de l'objet. Les trois types de caractéristiques étudiées au cours du projet sont: les coefficients d'harmonique circulaire (CHC), les sections en pointe (SP) et les descripteurs de Fourier (DF). La méthode basée sur la corrélation qui a été étudiée au cours du projet est la méthode DCCF (Distance Classification Correlation Filters). Cette dernière méthode permet d'aborder le problème de la classification des objets selon deux classes et, par le fait même, peut être généralisée au cas multiclasse. La méthode DCCF fait intervenir des procédés d'optimisation pour regrouper sur une même paire de filtres, l'information contenue dans plusieurs objets ou encore sur plusieurs vues d'une paire d'objets. Les filtres DCCF sont utilisés pour extraire un couple de paramètres de distance servant de critère de classification.

Les représentations basées sur les CHC, les SP et les DF sont toutes trois indépendantes de la taille apparente de l'objet. De plus, les signatures CHC et DF ne sont pas dépendantes de l'orientation (rotation dans le plan) de l'objet, tandis que la signature DF de même que la signature SP sont intrinsèquement indépendantes de la position de l'objet. Ces invariances permettent de réduire la complexité du problème de la clasification des signatures. Les méthodes d'extraction des caractéristiques, contrairement aux méthodes basées sur la corrélation, impliquent des opérations de préparation des images. Celles-ci sont destinées à éliminer l'arrière scène de l'image pour ne garder que l'objet d'intérêt ou l'une de ses parties

(le contour par exemple). Quelques approches ont été étudiées pour réaliser la préparation des images.

Les travaux effectués dans le cadre de ce projet ont été exécutés par l'INO et une équipe de l'université Laval, constituée de deux professeurs, d'un étudiant au doctorat et d'un autre à la maîtrise. Les études sur les caractéristiques SP et DF ont été effectuées par l'équipe de l'université Laval, tandis que l'INO a réalisé celles sur les caractéristiques CHC et la méthode DCCF. L'INO avait la tâche de réaliser le modèle d'essai, alors que l'implication de l'équipe de l'université Laval se limitait à faire des simulations numériques uniquement. Parmi toutes les méthodes étudiées, la méthode des CHC est celle qui a été retenue en vue de la réalisation du modèle d'essai. Étant donné la grande complexité du problème de RAC générale et les limites quant aux ressources disponibles, un problème simple a été considéré pour le modèle d'essai. Le problème considéré pour le modèle d'essai, consistait en la reconnaissance de modèles réduits sur arrière scène uniforme, et la capture des images se faisait avec une caméra vidéo sensible dans le domaine du visible. Pour ce problème, la préparation des images est peu complexe. Cette simplification à permis de consacrer la majorité des ressources à la réalisation du montage optique destiné à faire le calcul des CHC.

La scène type utilisée pour les simulations effectuées sur les méthodes SP et DF consistait en des véhicules militaires montés sur une plate-forme rotative et observés par une caméra IR. Dans la méthode SP, la segmentation des images consistait seulement à éliminer certains éléments de la scène et a été effectuée par un opérateur humain. Par contre, dans la méthode DF, les algorithmes utilisés pour la préparation des images étaient beaucoup plus élaborés. Puisque les caractéristiques DF sont extraites sur le contour fermé de la silhouette des objets, la préparation des images doit être effectuée avec soins. Ainsi, un filtre médian à été appliqué trois fois de suite sur chaque image dans le but de réduire le bruit, et l'extraction des contours des objets a été effectué à l'aide de la transformée en ondelettes. En raison du faible contraste des images, l'extraction des contours n'a pu être effectuée de façon entièrement automatique et une intervention humaine a été nécessaire lors les dernières étapes. La scène des véhicules militaires sur plate-forme rotative observée dans l'IR a également été utilisée pour des simulations sur la méthode des CHC et des DCCF. Dans ces derniers cas, la segmentation a été effectuée par un opérateur humain et consistait seulement à éliminer certaines portions de l'arrière scène.

L'étude de la méthode DCCF a été abandonnée à peu près au milieu du projet, en raison des difficultés de mise en oeuvre. Dans les contextes décrit dans le paragraphe précédent, les simulations pour les autres méthodes ont donné de bons résultats. Des taux d'erreur faibles ont été obtenus, toutefois il faut préciser que les résultats ont été obtenus sur un nombre restreint d'images et qu'une partie seulement des distorsions géométriques ont été considérées.

La mise en oeuvre optique de l'extraction des CHC a impliqué un effort important. Même au prix de tous les efforts consacrés, il n'a pas été possible de compléter le modèle d'essai prévu faute de temps. Ceci s'explique par le fait que l'infrastructure informatique avait été conçue pour fonctionner avec un montage optique qui devait permettre le calcul simultané de toutes les caractéristiques CHC. Toutefois ce montage optique, basé sur un réseau de diffraction, n'a pas donné les résultats escomptés et une solution de remplacement n'a été trouvée qu'à la fin du projet. Pour ce second montage optique, l'évaluation d'une signature CHC requière plusieurs cycles vidéo. Ceci impliquait des modifications substantielles aux différents logiciels déjà crées. Toutefois, il faut préciser qu'une bonne partie des éléments du modèle d'essai a été complétée et que le dernier montage a permis de faire l'extraction optique de signatures CHC. Ce même montage a également permis l'extraction de caractéristiques *CHC angulaires* qui sont une variante des caractéristiques CHC. Ces résultats ont permis des opérations de reconnaissance dans les contextes décrit précédemment. Une version miniaturisée du second montage a été proposée, celle-ci permettrait probablement de faire l'extraction de plusieurs caractéristiques CHC simultanément.

CHAPITRE 1

Synthèse

Ce rapport fait état des travaux qui ont été réalisés durant la seconde partie du projet commandité par le Centre de Recherche pour la Défense de Valcartier (CRDV) et intitulé : *Programme de recherche sur les réseaux de neurones optiques appliqués à la reconnaissance de cibles*. Le projet porte sur le développement d'algorithmes et d'un modèle d'essai (MIL-E-5400 voir annexe 5) destinés à illustrer les capacités des réseaux de neurones et des technologies du traitement optique de l'information pour la reconnaissance automatique de cibles (RAC). Le présent chapitre s'adresse aux lecteurs qui veulent prendre connaissance du projet dans son ensemble sans trop entrer dans les détails techniques. Ceux-ci sont fournis dans les chapitres suivants de même que dans le rapport de mi-projet dont le présent rapport fait suite.

1.1 Introduction

La reconnaissance des véhicules et autres équipements présents sur un théâtre d'opération militaire est une tâche de grande importance. Celle-ci est toutefois fort complexe en raison notamment de la grande diversité des aspects que peuvent prendre les objets. Généralement la reconnaissance doit se faire à partir d'informations partielles. Ceci est particulièrement vrai dans le cas où l'information est recueillie uniquement avec un système d'imagerie. Pour obtenir une description acceptable d'un objet non planaire il faut généralement un nombre relativement important d'images prises sous différents points de vue. La reconnaissance d'objets dans une scène imagée est un problème fort complexe mais certes pas insoluble. Comme argument pour appuyer cette affirmation on peut invoquer l'existence du système humain (et des animaux) qui réalise cette opération avec un bon degré de succès même dans des contextes difficiles.

Les remarquables capacités de l'observateur humain à reconnaître les objets à partir d'information 2D uniquement lui sont conférées par son 'système' (le cerveau) d'interprétation des images très performant. Les capacités de la perception visuelle humaine sont difficilement reproductibles par une machine, toutefois les réseaux de neurones artificiels (RNA) offrent des perspectives intéressantes. À l'instar des systèmes nerveux biologiques, les RNA sont des architectures de traitement constituées d'un nombre important d'éléments simples qui sont massivement interconnectés les uns aux autres. Cette organisation distribuée du traitement confère aux RNA des propriétés similaires à celles de leurs homologues biologiques. Entre autres les RNA offrent la possibilité d'assimiler l'information de façon plus ou moins autonome en une représentation compacte. Ceci se fait lors d'un processus appelé entraînement, au cours duquel les poids des interconnexions sont itérativement modifiés de manière à produire le comportement désiré. Les RNA sont particulièrement intéressants pour la modélisation implicite de situations complexes pour lesquelles il n'existe pas de modèles explicites adéquats comme c'est le cas en reconnaissance d'objet 3D à partir d'information 2D (image).

La reconnaissance de formes est l'un des sujets parmi les plus étudiés en traitement optique de l'information (TOI). L'optique s'impose tout naturellement pour le traitement d'informations bidimensionnelles (2D) en raison de la capacité inhérente des lentilles à effectuer la transformée de Fourier, laquelle opération intervient dans plusieurs méthodes de reconnaissance de formes. Le parallélisme massif de l'optique offre également des perspectives intéressantes pour l'extraction des propriétés caractérisant les images. Ceci permet la génération de représentations compactes dans lesquelles l'information morphologique (information sur les formes) est convertie en un ensemble de paramètres numériques. De plus l'optique permet de propager l'information en espace libre. Cette propriété représente un attraits intéressants pour la mise en oeuvre des interconnexions nécessaires à la réalisation de RNA.

Les capacités de l'optique à effectuer du traitement en parallèle de même que les perspectives de la mise en oeuvre de RNA à l'aide de l'optique sont les principales motivations qui sont à l'origine du projet décrit dans ce rapport. Le projet se veut une investigation non exhaustive des possibilités des technologies optiques pour la mise en oeuvre d'algorithmes classiques et/ou neuronaux de reconnaissance d'objets à partir d'images. Plus

spécifiquement, les travaux s'insèrent dans une perspective militaire et les objets à reconnaître correspondent à des cibles potentielles.

Pour l'exécution d'une partie de ce projet, l'institut national d'optique (INO) a fait appel aux services de l'Université Laval. La contribution de l'Université Laval au projet a été limitée au développement et à la simulation de certaines méthodes de reconnaissance automatique de cibles (RAC). La contribution de l'INO au projet consistait en le développement et la simulation d'autres méthodes de RAC ainsi qu'à la mise en oeuvre du modèle d'essai.

1.2 Problématique et approches

1.2.1 Énoncé du problème de RAC

Le problème de reconnaissance automatique de cibles d'intérêt pour le CRDV est illustré à la figure 1.1. Il s'agit d'une scène perçue par une caméra qui opère ici dans la partie infrarouge (IR) du spectre. Le problème consiste à déceler la présence des cibles et à déterminer à quel type (ex : char, camion, etc) de véhicules appartient chacune des cibles, cela en temps réel c'est-à-dire à une cadence comparable à celle de l'acquisition des images (30 images/ seconde). Le mode d'acquisition de l'information n'est pas parfaitement adapté au type de problème de reconnaissance puisque la scène recueillie est bidimensionelle alors que l'information requise est répartie dans un volume (3-D). De cela découle en partie le problème des distorsions géométriques, à savoir qu'un même objet peut prendre un grand nombre d'aspects différents selon l'information accessible (projections) par la caméra 2D.

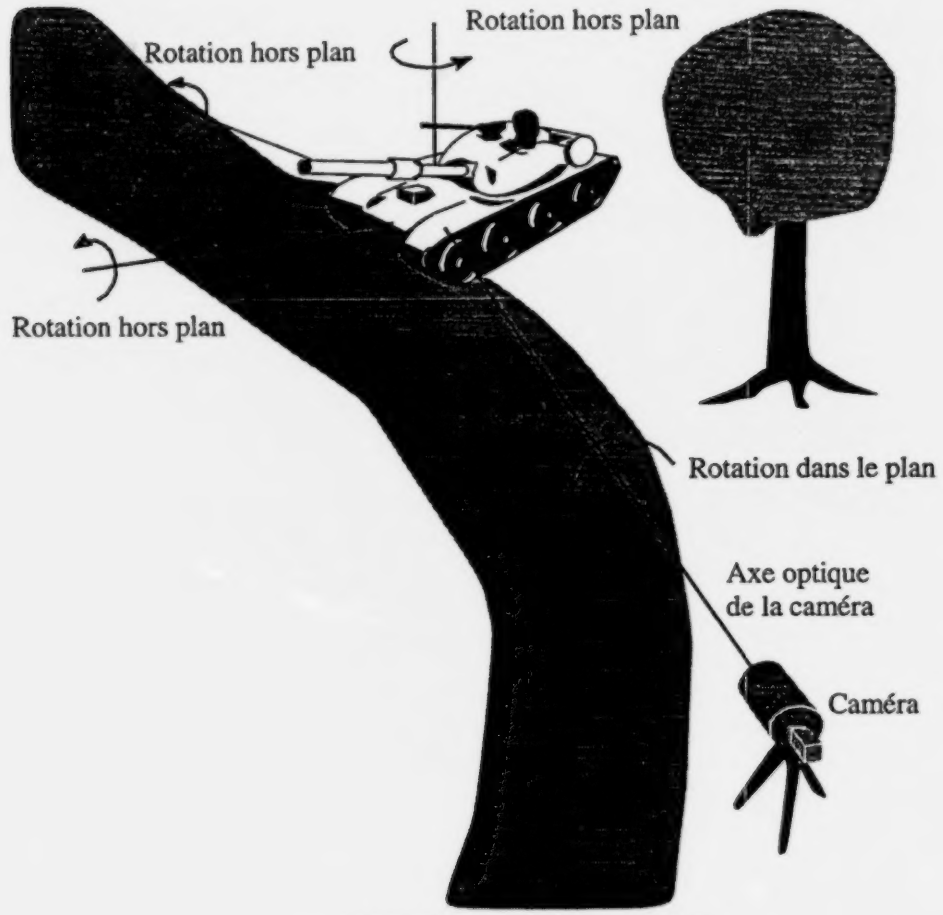


FIGURE 1.1 : Problème de RAC

Les distorsions géométriques sont définies comme étant les changements d'aspect d'un objet selon son orientation et sa position relativement à la caméra. On considère ici trois types de distorsions géométriques, celles-ci sont illustrées sur la figure 1.1. Il y a tout d'abord le *changement d'échelle* qui se produit lorsque la distance entre la caméra et la cible change ou encore lorsque la focale de l'optique de la caméra est modifiée. La *rotation dans le plan* est le type de distorsion géométrique qui s'observe lorsque l'objet subit une rotation autour de l'axe optique de la caméra. Le dernier type de distorsion considéré est la *rotation hors plan.* Celle-ci correspond à une rotation de l'objet autour de l'un des axes perpendiculaires à l'axe

optique de la caméra. La rotation hors plan est la distorsion géométrique qui cause le plus de difficultés pour la reconnaissance.

1.2.2 Éléments de solution

Les RNA et le TOI ont été proposées comme moyens pour aborder le problème de reconnaissance de formes. L'idée d'utiliser l'optique pour faire la mise en oeuvre de RNA a été étudiée au cours de la première partie du projet. Toutefois, après analyse cette idée a été mise de côté puisque les processeurs numériques actuellement disponibles sont suffisamment puissants pour permettre la mise en oeuvre des RNA d'architectures simples. Les composants optiques n'ont pas encore la maturité requise pour permettre la réalisation optique des architectures de RNA complexes. Certaines architectures de RNA (le Néocognitron par exemple) permettent, dans une certaine mesure, la reconnaissance automatique d'objets directement sur les images. Ces architectures de RNA sont complexes et impliquent une charge de traitement numérique trop importante pour pouvoir opérer en temps réel. La stratégie adoptée dans ce projet vise à réduire la dimensionalité du problème en utilisant des méthodes conventionnelles, pour ensuite faire intervenir des RNA d'architecture simple pouvant être mis en oeuvre sur un ordinateur de capacité modeste. Les méthodes conventionnelles de RAC que l'on retrouve fréquemment dans les publications spécialisées font soit appel à la corrélation, ou encore impliquent la représentation des objets par un ensemble de caractéristiques morphologiques. Dans les deux cas, il y a possibilité de faire intervenir l'optique pour la mise en oeuvre.

L'approche dite à extraction des caractéristiques, consiste à associer à chacune des portions de la scène correspondant à un objet, un nombre fini de paramètres représentatifs de la morphologie de l'objet. L'ensemble de ces paramètres constitue une ***signature*** qui sert à représenter l'objet. La classification des objets se fait donc à partir de sa signature à l'aide d'un classificateur conventionnel ou encore à l'aide d'un RNA d'architecture relativement simple. Cette approche implique toutefois une étape de préparation de la scène que l'on peut diviser en trois parties successives: la réduction du bruit, la détection et la segmentation. La détection consiste en l'identification des régions (d'intérêt) pouvant contenir une cible. La segmentation quant à elle, permet pour chacune de ces régions d'intérêt, d'isoler de l'arrière scène la portion assimilable à l'objet présumé. Le succès de la classification des objets à partir de la représentation mathématique constituée de l'ensemble ordonné des caractéristiques

dépend grandement de l'aptitude de ces dernières à bien représenter les objets malgré les changements d'aspects. À notre connaissance il n'existe pas de caractéristiques morphologiques qui permettent de constituer une représentation qui soit invariante sous l'ensemble des changements d'aspects.

L'approche dite par corrélation ne nécessite pas en principe les étapes de prétraitement requises par l'approche à extraction des caractéristiques. La corrélation est une mesure de la similitude entre une figure de référence et les différentes portions d'une image. Le résultat de la corrélation est une image appelée *plan de corrélation*. En général, la présence d'un objet similaire à la figure de référence est révélée par l'apparition de marqueurs (ex. une tache très brillante) dans le plan de corrélation. La corrélation se prête bien à la reconnaissance d'une figure (objet 2D) spécifique, toutefois la reconnaissance d'objets se manifestant sous différents aspects n'est pas chose facile. Les changements d'aspects d'un objet peuvent être pris en compte en créant une figure de référence composite. Celle-ci est en quelque sorte une superposition de plusieurs vues de l'objet représentatives d'une partie de l'ensemble des aspects selon lesquels l'objet peut se manifester. Pour tenir compte de la plage toute entière des changements d'aspects, il faut en général un nombre important de ces figures de référence (***filtres***).

1.2.3 Description des méthodes étudiées

On retrouve dans les publications un grand nombre de méthodes qui ont été proposées pour aborder le problème de la RAC. L'une des toutes premières tâches dans ce projet a consisté à faire un inventaire non exhaustif des méthodes existantes pour ensuite choisir celles qui sont les plus adaptées ou les mieux adaptables à notre contexte de RAC. Au cours de la première partie du projet, trois méthodes de caractérisation morphologiques, ainsi que la méthode de corrélation DCCF ont été étudiés. Le choix des méthodes étudiées faisait suite à une étude bibliographique effectuée au cours des premiers mois du projet. Les types de caractéristiques étudiées sont les coefficients d'harmoniques circulaires (CHC), les sections en pointe et les descripteurs de Fourier sur les contours. La méthode des CHC et celle des DCCF ont été étudiées par l'INO, tandis que les méthodes des sections en pointe et des descripteurs de Fourier ont été étudiées par l'équipes des professeurs Arsenault et Sheng de l'université Laval. Les types de caractéristiques étudiées et la méthode DCCF sont brièvement décrites aux pages suivantes.

1.2.3.1 Coefficients d'harmoniques circulaires (CHC)

Les ***coefficients d'harmoniques circulaires*** (CHC) sont expliqués schématiquement à la figure 1.2. On considère un objet sur un fond noir uniforme. L'intensité de l'objet peut être représentée par une fonction $f(r,\theta)$ exprimée en terme des coordonnées polaires r et θ. Pour chacune des valeurs r_0 de la coordonnée r, l'expression :

$$f(\theta;r_0) \equiv f(r,\theta)\,|_{r=r_0}$$

décrit une fonction de la variable angulaire θ uniquement. Puisque $f(\theta;r_0)$ est périodique de période 2π, elle peut être décomposée en une série de Fourier [1]:

$$f(\theta;r_0) = \sum_{m=-\infty}^{\infty} a_m e^{im\theta}$$

$$\text{avec } a_m = \frac{1}{2\pi}\int_0^{2\pi} f(r,\theta)\,e^{im\theta}\,d\theta\Big|_{r=r_0} = a_m(r)\,|_{r=r_0}.$$

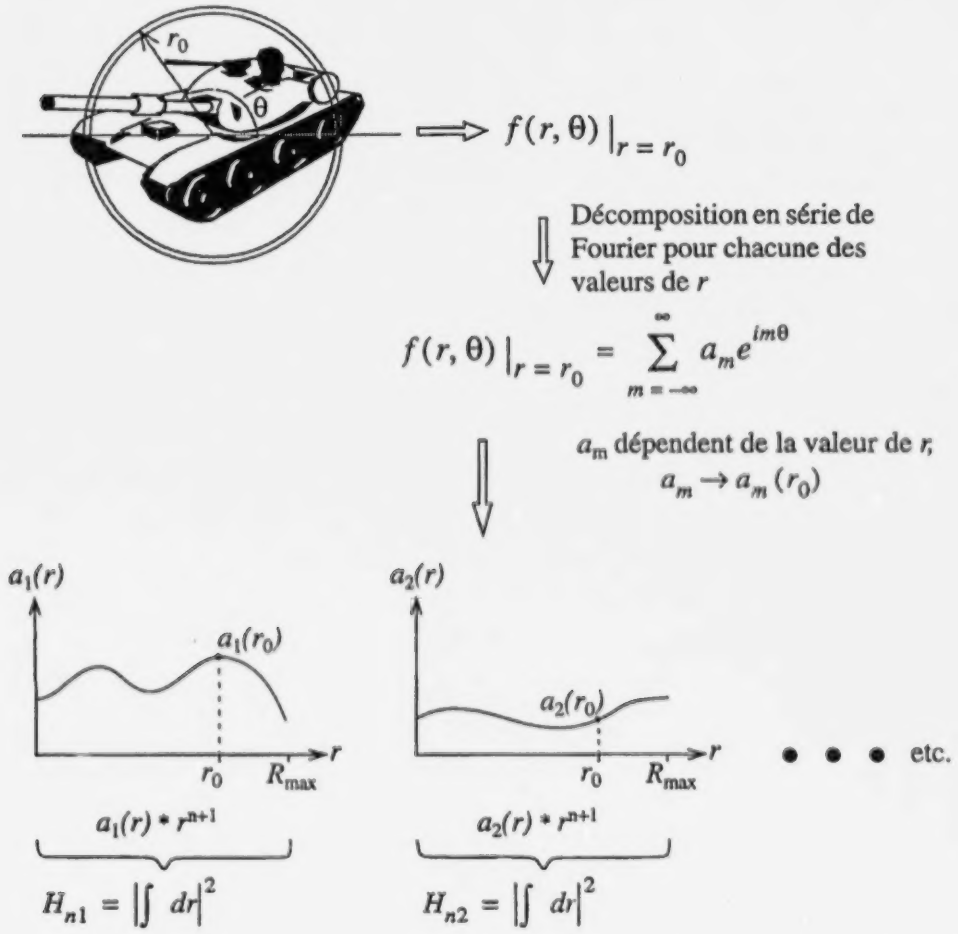


FIGURE 1.2 : Explication schématique des coefficients d'harmoniques circulaires (CHC).

On retrouve un coefficient a_m, quelque soit l'entier m et pour toute valeur r_0 de la coordonnée r. Ainsi l'ensemble des a_m constitue une fonction $a_m(r)$ qui dépend de la variable r. Le coefficient d'harmonique circulaire H_{nm} d'ordre n,m est défini comme étant le module carré de la somme intégrale du produit de la fonction $a_m(r)$ avec la fonction de pondération radiale r^{n+1}:

$$H_{nm} \equiv \left| \frac{1}{\pi} \int_0^{R_{max}} a_m(r)\, r^{n+1} dr \right|^2 = \left| \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} f(r,\theta)\, e^{-im\theta} r^{n+1} d\theta dr \right|^2 \qquad (0.1)$$

Les CHC sont indépendants de l'orientation de l'objet et, moyennant une normalisation adéquate, ils sont également indépendants de l'échelle de l'objet. Toutefois, ils dépendent de la position de l'objet relativement à l'origine des axes de coordonnées. Ce problème peut se résoudre par un recentrage de l'objet ou encore en calculant les CHC non pas sur l'objet lui-même mais plutôt sur le module de sa transformée de Fourier (TF). Le module de la TF est une fonction indépendante de la position de l'objet et c'est cette dernière solution qui a tout d'abord été retenue. Les CHC d'un objet ont alors été définis comme étant l'intégrale 0.1 évaluée sur le module carré de la TF de l'objet. La mise en oeuvre optique du calcul des CHC peut se faire de deux façons. La première consiste à utiliser la transformée de Fourier. En effet l'intégrale de la relation 0.1 correspond à la valeur à l'origine de la TF de la fonction $r^n f(r,\theta)\, e^{-im\theta}$. La seconde façon de faire implique la décomposition de la relation (0.1) en ses parties réelle et imaginaire. En poussant d'avantage le développement de façon à obtenir des arguments à valeurs positives seulement, on obtient la relation suivante :

$$H_{nm} = 4\,(C_{nm} - C_{n0})^2 + 4\,(S_{nm} - C_{n0})^2 \qquad (0.2)$$

où C'_{nm} et S'_{nm} sont les sommes intégrales en coordonnées polaire des fonctions:

$$r^n f(r,\theta) \left[\frac{1 + \cos(m\theta)}{2\pi} \right] \qquad (0.3)$$

et

$$r^n f(r,\theta) \left[\frac{1 + \sin(m\theta)}{2\pi} \right] \qquad (0.4)$$

respectivement. La mise en oeuvre optique peut donc se faire en projetant l'image d'entrée $f(r,\theta)$ sur différents masques d'amplitudes $r^n[1+\cos(m\theta)]/2$ et $r^n[1+\sin(m\theta)]/2$ et de recueillir la lumière transmise par chacun des masques sur un détecteur. Ces masques sont ce que l'on

appelle des ***rosettes*** avec fonction de pondération radiale. La rosette en $1+ cos(m\theta)$ est dite complémentaire de celle en $1+ sin(m\theta)$ de même ordre harmonique m et vice versa.

1.2.3.2 Sections en pointe [2]

Les sections en pointe sont des caractéristiques que l'on calcule sur le module de la TF de l'objet à classifier. Ils correspondent à la somme de l'énergie (module carré de l'amplitude) contenue dans les différentes portions en forme de pointe. Ceci est illustré à la figure 1.3. La normalisation de la signature permet d'obtenir l'invariance sous échelle, et puisque les caractéristiques sont extraites sur le module de la TF, la signature est intrinsèquement indépendante de la position de l'objet dans la scène.

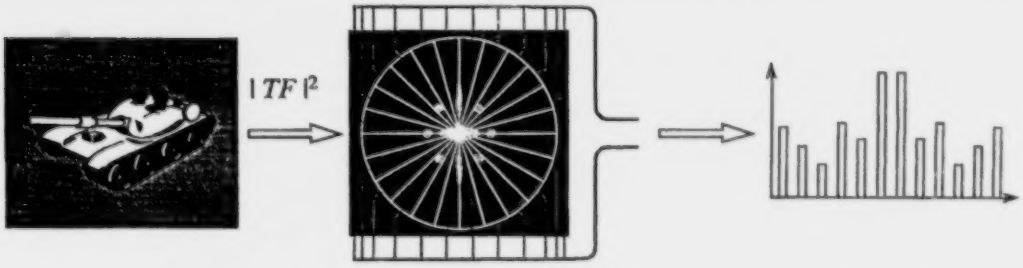


FIGURE 1.3 : Les caractéristiques des sections en pointe

1.2.3.3 Descripteurs de Fourier [3]

La méthode des descripteurs de Fourier fait intervenir le contour fermé de la silhouette de l'objet. Un point de départ est choisi pour devenir l'origine de la coordonnée curviligne l qui court le long du contour. Les angles de la tangente à la courbe mesurés relativement à la tangente à la courbe au point de départ constituent une fonction $\phi(l)$ qui prend des valeurs de 0 à 2π. La nouvelle fonction $\phi^*(t) = \phi(Lt/2\pi) + t$, où L est la longueur totale du contours et t est un paramètre qui varie de 0 à 2π, est une fonction périodique de période 2π. La fonction $\phi^*(t)$ est développée en série de Fourier et le module $|A_k|$ des coefficients de la série constitue la signature des descripteurs de Fourier. Cette signature est invariante sous échelle et sous rotation dans le plan et ne dépend pas de la position de l'objet. Le calcul des descripteurs de Fourier est illustré à la figure 1.4

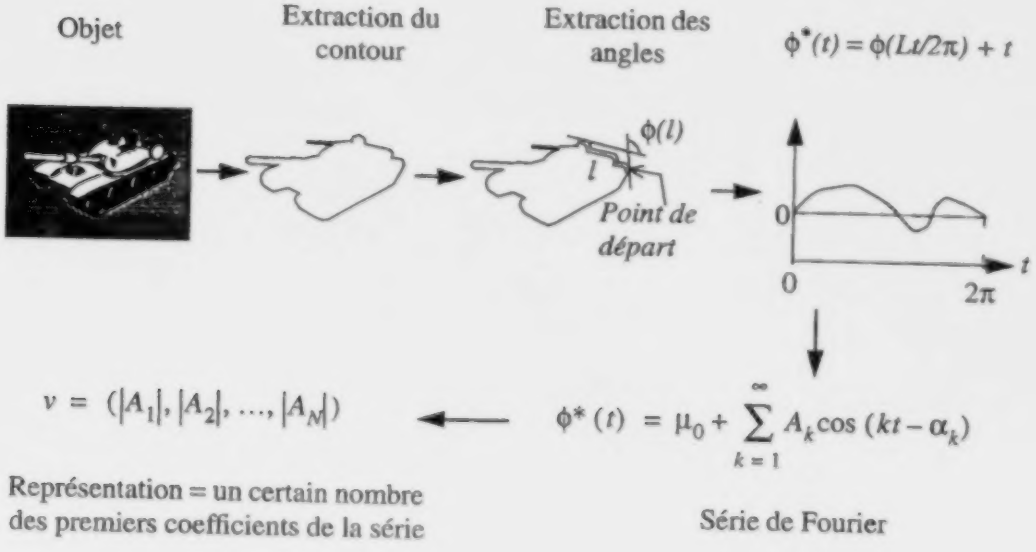


FIGURE 1.4 : Descripteurs de Fourier

1.2.3.4 Méthode des DCCF [4]

En ce qui concerne la reconnaissance par corrélation, c'est la méthode DCCF (Distance Classification Correlation Filters) qui a été étudiée. Cette méthode est conçue pour traiter les problèmes pour lesquels les objets à reconnaître sont regroupés en deux classes. Puisque le cas multiclasses peut être ramené à une série de problèmes à deux classes, la méthode DCCF peut donc être généralisée au cas multiclasses. La méthode DCCF consiste à créer une figure de référence, ou filtre, qui fait en sorte que la corrélation prenne des valeurs au centre (point du plan de corrélation correspondant au centre de l'objet) qui soient les plus semblables possible pour les objets d'une même classe tout en étant les plus différentes possibles d'une classe à l'autre. La méthode DCCF fait appel à des techniques d'optimisation pour obtenir ces propriétés tout en minimisant l'énergie globale du plan de corrélation. À partir du filtre DCCF, deux autres figures de référence sont calculées, soit une pour chacune des deux classes. La fonction de référence associée à la première classe prend des valeurs élevées au centre uniquement lorsque l'image comprend un objet appartenant à la première classe et vice versa pour la seconde figure de référence. Ainsi, si l'image de la scène comprend un objet appartenant à l'une ou l'autre des classes, l'une des deux figures de référence doit produire un plan de corrélation ayant une valeur élevée au point correspondant à la position de l'objet dans la scène tandis que pour l'autre figure de référence la corrélation devrait prendre une valeur plutôt faible en ce point.

La méthode DCCF implique un post-traitement relativement important puisque les deux plans de corrélation doivent être analysés pour détecter les couples de pic/anti-pic. Au couple de valeurs du pic et de l'anti-pic associée à un objet donnée correspond, après une certaine transformation, un point dans l'espace des caractéristiques (espace 2D dans le cas présent). En principe, dans cet espace, les points correspondant aux objets de la première classe doivent être regroupés en un amas disjoint de celui des points correspondant aux objets de l'autre classe. La classification des objets peut se faire à l'aide d'un réseau de neurones ou d'un classificateur conventionnel, par le biais du couple des valeurs dérivées des valeurs pic/anti-pic. La méthode DCCF est illustrée à la figure 1.5.

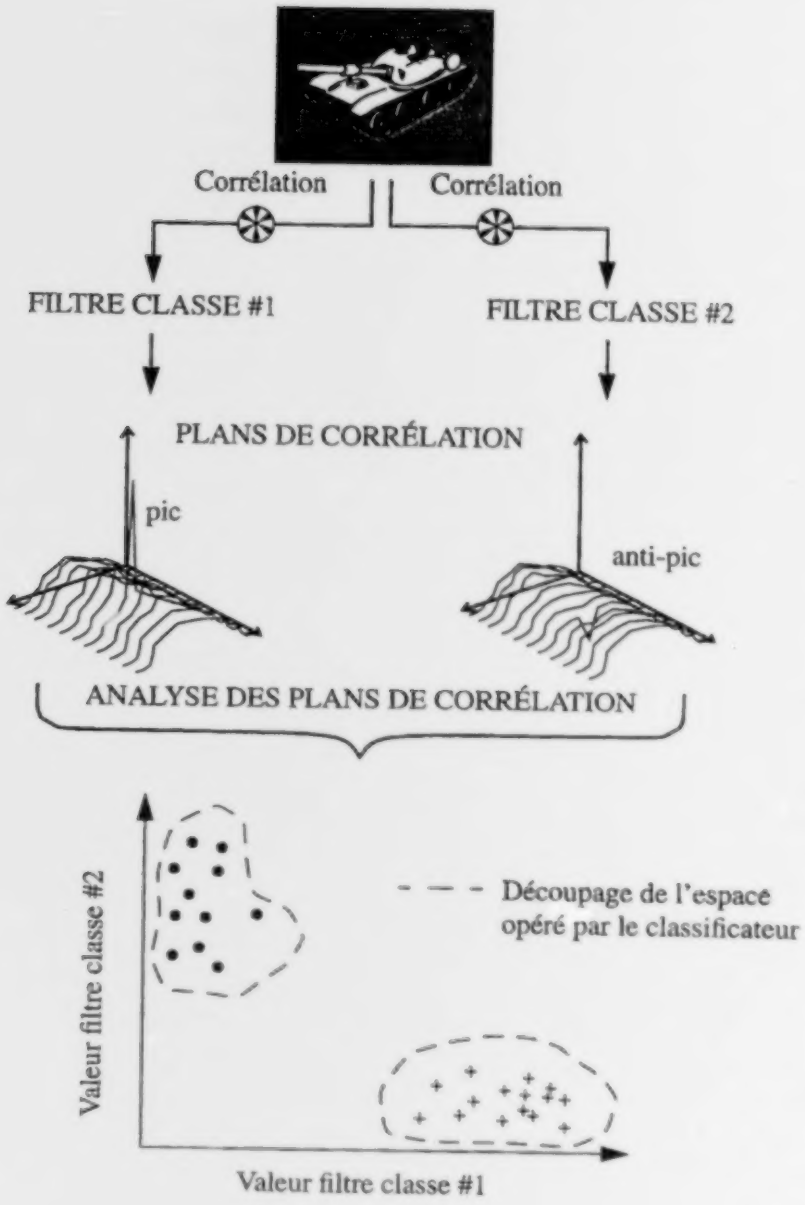


FIGURE 1.5 : Méthode DCCF

1.3 Sommaire de la première partie du projet

Les travaux exécutés dans le cadre de la première partie du projet sont décris dans le rapport de mi-projet. Cette section se veut un résumé de l'information contenue dans ce dernier rapport, le lecteur désireux de prendre connaissance des détails est invité à consulter le rapport de mi-projet.

1.3.1 Simulations numériques

Des simulations ont été effectuées lors de la première partie du projet pour évaluer les performances de la méthode des CHC. Les simulations ont portées dans un premier temps sur le cas plus simple pour lequel on considère uniquement les distorsions dans le plan. Ce dernier cas impliquant seulement des changements d'échelle et des rotations dans le plan est appellé le ***problème 2D***. Le cas plus complexe, pour lequel les trois types de distorsions géométriques sont considérées est appellé ***problème 3D***. Pour les simulations du problème 2D, des images de différents modèles d'aéronefs sur fond noir uniforme ont été utilisés. Ces simulations ont démontré que les rotations dans le plan et les changements d'échelle affectent très peu la valeur des CHC, cela malgré la pixellisation et la résolution relativement faible des images et des rosettes. Un réseau de neurones de type perceptron a été entraîné à classer correctement les signatures CHC. Les images utilisées pour les simulations du problème 3D sont les images IR de chars en rotation (hors plan) sur une plate-forme tournante (images fournies par le CRDV). Un nombre important des images de la séquence vidéo a été numérisé et traité par un opérateur humain afin d'éliminer l'arrière scène. Pour effectuer la classification des signatures CHC extraites sur ces images, une série de perceptrons organisés en une structure hiérarchique a été utilisée. Le classificateur neuronal effectue la classification par étape, en statuant dans un premier temps sur l'orientation grossière (avant, arrière, cotés) de l'objet. Selon le résultat de cet première étape de classification, la signature CHC est présentée à l'un des perceptrons du dernier étage de la hiérarchie. Chacun des perceptrons du dernier étage prend en charge la classification des objets pour une portion de la plage globale de la rotation hors plan (selon l'axe de la plate-forme tournante). En utilisant ce type de classificateur neuronal, il a été possible de classer un nombre restreint de types de chars avec relativement peu d'erreurs.

L'aptitude de la méthode DCCF a également fait l'objet d'une évaluation par le biais de simulations numériques. Celles-ci ont été uniquement effectuées sur le problème 3D en utilisant les images IR des chars en rotation sur plate-forme tournante. Les tests ont porté sur un cas à deux classes constituées de différentes versions tournées de deux types de chars. Il a été possible de calculer une fonction de référence permettant la séparation des classes dans l'espace de classification. Toutefois, même en comptant sur les capacités de 'généralisation' du filtre DCCF, les simulations ont mis en évidence le fait que la plage de changements d'aspect qui peut être 'enregistrée' par une fonction de référence est relativement limitée.

La partie de l'étude exécutée par l'université Laval n'a pu débuter que plusieurs mois après le début du projet. Aussi les travaux portant sur les sections en pointes et les descripteurs de Fourier n'étaient pas très avancés au moment de la rédaction du rapport de la première partie du projet. Toutefois, les résultats préliminaires sur la méthode des sections en pointes avaient été qualifiés de prometteurs par l'équipe de l'université Laval.

1.3.2 Montages optiques

Pour explorer la possibilité de faire la mise en oeuvre optique de l'extraction des CHC, deux systèmes optiques rudimentaires ont été assemblés. Un premier montage à été assemblé pour faire l'extraction des CHC par le biais de la transformée de Fourier optique. Ce premier montage n'a pas donné de résultats valables en raison principalement de la présence du bruit sur la TF optique. Étant donné les difficultés rencontrées avec le premier montage, un second montage a été réalisé. Ce dernier était basé sur la méthode des rosettes en $1+cos$ et $1+sin$. La TF de l'objet affichée sur le premier ACL est projetée sur un second ACL lequel affiche à tour de rôle les différentes rosettes. L'éclairement est assuré par un laser He-Ne et l'énergie transmise par les deux ACL est recueillie sur un photodétecteur. Un ordinateur PC commande l'affichage des images sur les deux ACL de même que l'acquisition du signal provenant du détecteur. Les tests optiques d'extraction des CHC effectués au cours de la première partie du projet ont portés sur le problème 2D uniquement. Les images qui ont été utilisées pour ces tests étaient des images de modèles réduits d'aéronefs vus du dessus sur fond uniforme. Deux classes ont été constituées, la première regroupait trois versions tournées (rotation dans le plan) de la navette spatiale de même que plusieurs versions de différentes tailles tandis que la seconde classe était constituée de 5 images de différents avions. Les signatures CHC ont été évaluées optiquement et un perceptron a pu être entraîné à classer les

aéronefs correctement selon les 2 classes. Ces résultats s'avéraient encourageant, toutefois ils ne démontraient pas l'invariance complète sous rotation des signatures. Il existait des variations significatives dans les signatures des trois différentes versions tournées de la navette.

Suite aux essais préliminaires sur l'extraction optique des CHC, deux nouveaux montages ont été proposés pour la suite du projet. Ces deux nouveaux montages ont pour point commun le parallélisme qui doit en principe permettre l'extraction de plusieurs valeurs de CHC simultanément. De plus, dans ces deux montages, les rosettes s'affichent sur des supports permanents plutôt que sur un ACL. Pour le premier des montages proposés, l'image de l'objet affichée sur un ACL est projetée sur un réseau de diffraction pour ainsi produire la duplication des images. Les faisceaux de lumière transmis par le réseau de diffraction passent ensuite dans une lentille, pour produire dans le plan focal de cette dernière plusieurs répliques de la TF de l'objet, lesquelles éclairent les différentes rosettes. Pour le second montage, la TF optique de l'objet est capturée par une caméra vidéo. L'image capturée est ensuite répliquée numériquement de façon à créer une mosaïque destinée à être affichée sur un ACL. Une fois affichée sur l'ACL, la mosaïque de TF est projetée sur une mosaïque de rosettes.

Des travaux préliminaires dans le but de faire la mise en oeuvre optique de la méthode DCCF ont été effectués au cours de la seconde moitié de la première partie du projet. Les tests effectués à l'aide d'un corrélateur optique monté sur une table optique n'ont pas donné des résultats concluants. Les filtres DCCF sont plutôt difficiles à réaliser optiquement, puisque l'afficheur à cristaux liquides[1] (ACL) utilisé pour l'affichage dynamique des filtres dans le plan de Fourier du corrélateur est limité quant aux niveaux de gris et au codage des phases réalisables. Pour réaliser optiquement la méthode DCCF il faut adapter les filtres de façon à tenir compte du nombre restreint de degrés de liberté offert par l'ACL. Malgré l'utilisation d'un algorithme d'optimisation permettant d'exploiter au mieux les possibilités de l'ACL, les résultats optiques obtenues sont dans l'ensemble peu satisfaisant. D'ailleurs, une partie seulement de la méthode DCCF a été réalisée, soit seulement la mise en oeuvre du filtre correspondant à la fonction de référence. Le reste de l'algorithmes permettant l'extraction de la valeur des couples pic/anti-pic n'a pu être mise en oeuvre optiquement. De plus, les plans

1. Ceci est le cas pour la grande majorité des ACL.

de corrélation obtenus optiquement avec le filtre de la fonction de référence DCCF contenaient beaucoup de bruit rendant ainsi difficile et hasardeux l'interprétation des plans de corrélation.

1.3.3 Traitement numérique

Les méthodes à extraction des caractéristiques impliquent différentes étapes de traitement. Ces étapes de préparation des images sont essentielles puisque les caractéristiques doivent être extraites sur les objets seuls, la présence de bruit et de l'arrière scène pouvant éventuellement influencer la valeur des caractéristiques extraites. Le type de traitement impliqué dans ces étapes de segmentation se prête bien à une mise en oeuvre numérique. L'INO dispose d'un processeur numérique parallèle de marque ***Datacube,*** capable d'effectuer des traitements élémentaires à cadence élevée. Disposant de cet appareil, et d'une certaine experience dans sa programmation, il a été décidé d'utiliser le Datacube pour la mise oeuvre des étapes de préparation des images.

Dans le contexte des images IR, la détection des cibles potentielles a été réalisée sur le Datacube en utilisant une simple opération de seuillage. Le résultat de l'opération est une image binaire (comportant deux niveaux de gris seulement) sur laquelle les seuls pixels à valeur non nulle sont ceux de l'image originale ayant des valeurs supérieures à celle du seuil. Puisqu'il s'agit d'images IR, l'opération de détection permet le repérage des zones chaudes dans la scène. La segmentation a été réalisée par filtrage de type ***Sobel*** [5], mis en oeuvre sur le Datacube. L'application du filtre de Sobel permet d'estimer le gradient des variations de l'intensité dans l'image. Un seuillage sur l'image du gradient fait en sorte que tous les pixels de l'image prennent la valeur nulle sauf ceux qui correspondent aux valeurs élevées du gradient où l'on retrouve généralement les contours des objets. Le résultat de la segmentation effectuée est donc une image binaire contenant en principe uniquement les contours de l'objet. L'application du filtre de Sobel se fait uniquement dans les régions d'intérêt identifiées lors de l'étape de la détection. L'exécution des algorithmes de détection et de segmentation réalisés sur le processeur Datacube maintient la cadence vidéo.

1.4 Sommaire de la seconde partie du projet

L'orientation de la seconde partie du projet découle des résultats obtenus lors de la première partie. En avril 1995, la méthode des CHC a été identifiée comme le choix le plus approprié. Il a été décidé que le modèle d'essai optique serait basé sur la méthode des CHC, compte tenu que les travaux sur les méthodes étudiées par l'université Laval étaient alors peu avancés et la méthode des DCCF s'avérait difficile à réaliser optiquement. Bien que les méthodes étudiées par l'équipe de l'université Laval ne soient pas utilisées pour la réalisation du modèle d'essai, elles n'en demeurent pas moins intéressantes et des efforts ont été consacrés pour compléter leurs évaluations. Ainsi, au cours de la seconde partie du projet, l'INO a travaillé sur la réalisation du modèle d'essai optique tout en améliorant ses méthodes algorithmiques alors que l'équipe de l'université Laval a poursuivi les travaux qu'elle avait entamés au cours de la première partie du projet. Les prochains paragraphes décrivent sommairement les travaux effectués au cours de la seconde partie du projet. Les détails sont fournis dans les chapitres subséquents du rapport.

1.4.1 Étude sur la méthode des sections en pointe et sur la méthode des descripteurs de Fourier

L'équipe de l'université Laval a poursuivi les travaux qu'elle avait débutés lors la première partie du projet. Ces travaux portaient sur les méthodes dites des sections en pointe et des descripteurs de Fourier. Les études sur ces méthodes se limitaient uniquement à de l'expérimentation numérique et n'impliquaient pas la réalisation de systèmes optiques. Les images utilisées pour faire les simulations numériques proviennent des séquences vidéo infrarouges fournies par le CRDV, sur lesquelles on retrouve des véhicules militaires montés sur une plate-forme tournante. Les prochaines sections décrivent brièvement les travaux effectués.

1.4.1.1 Méthode des sections en pointe

Il s'agit d'une méthode de classification des objets basée sur les caractéristiques des sections en pointe décrite à la section 1.2.3.2. La figure 1.6 illustre les étapes impliquées dans la méthode. La première étape est évidemment la segmentation de l'objet. Le module carré de la TF (spectre de puissance) de l'objet est ensuite calculé. Les valeurs de l'énergie sur les sections en pointe sont déterminées et la signature obtenue est normalisée relativement à

l'energie totale du spectre de Fourier. La classification des objets à partir de leur signature se fait à l'aide de la méthode des trajectoires dans l'espace des caractéristiques. La trajectoire associée à un objet donnée se fait à partir d'un *ensemble témoin* constitué d'un nombre restreint de vues représentatives de l'ensemble de toutes les vues possibles. Les signatures d'un objet de l'ensemble témoin sont des points dans l'espace multidimensionnel des caractéristiques. Ces points sont reliés les uns aux autres selon un ordre donné, ce qui crée une trajectoire dans l'espace des caractéristiques. On compte autant de trajectoires qu'il y a d'objets. La classification d'un objet se fait à l'aide d'un critère de proximité. La plus petite des distances euclidiennes qui séparent le point correspondant à l'objet à chacune des trajectoires, détermine la classe à laquelle l'objet est associé.

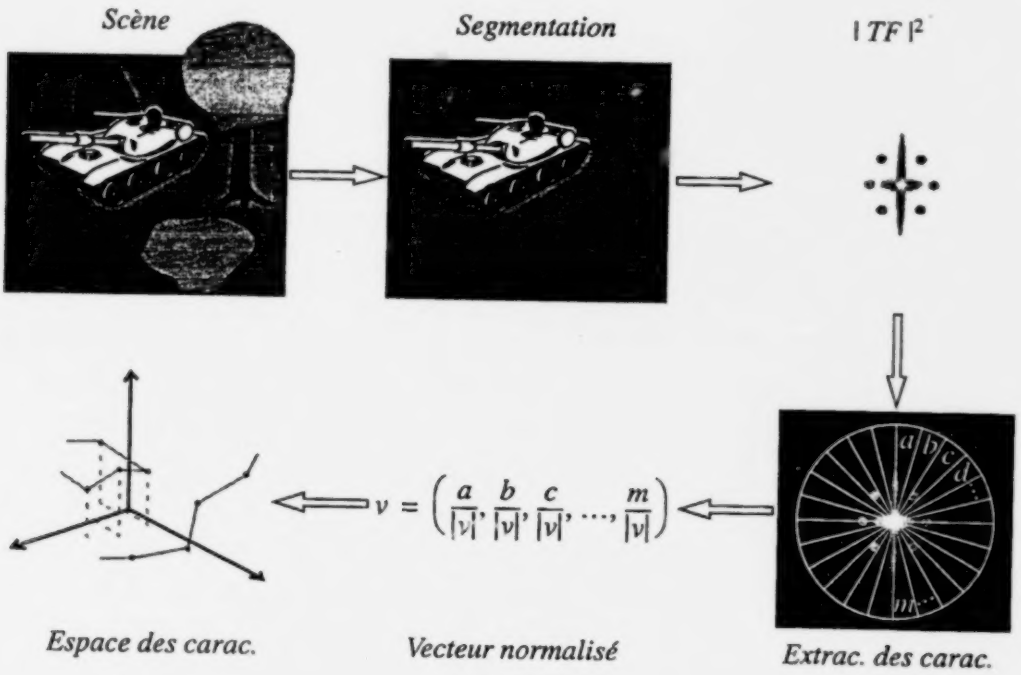


FIGURE 1.6 : Méthode des sections en pointe

Les simulations numériques ont été effectuées sur six types de véhicules. Ceux-ci ont été regroupés pour former deux classes. Des tests ont été effectués pour évaluer les performances de la méthode pour la rotation hors axe autour de l'axe vertical et pour les

l'energie totale du spectre de Fourier. La classification des objets à partir de leur signature se fait à l'aide de la méthode des trajectoires dans l'espace des caractéristiques. La trajectoire associée à un objet donnée se fait à partir d'un *ensemble témoin* constitué d'un nombre restreint de vues représentatives de l'ensemble de toutes les vues possibles. Les signatures d'un objet de l'ensemble témoin sont des points dans l'espace multidimensionnel des caractéristiques. Ces points sont reliés les uns aux autres selon un ordre donné, ce qui crée une trajectoire dans l'espace des caractéristiques. On compte autant de trajectoires qu'il y a d'objets. La classification d'un objet se fait à l'aide d'un critère de proximité. La plus petite des distances euclidiennes qui séparent le point correspondant à l'objet à chacune des trajectoires, détermine la classe à laquelle l'objet est associé.

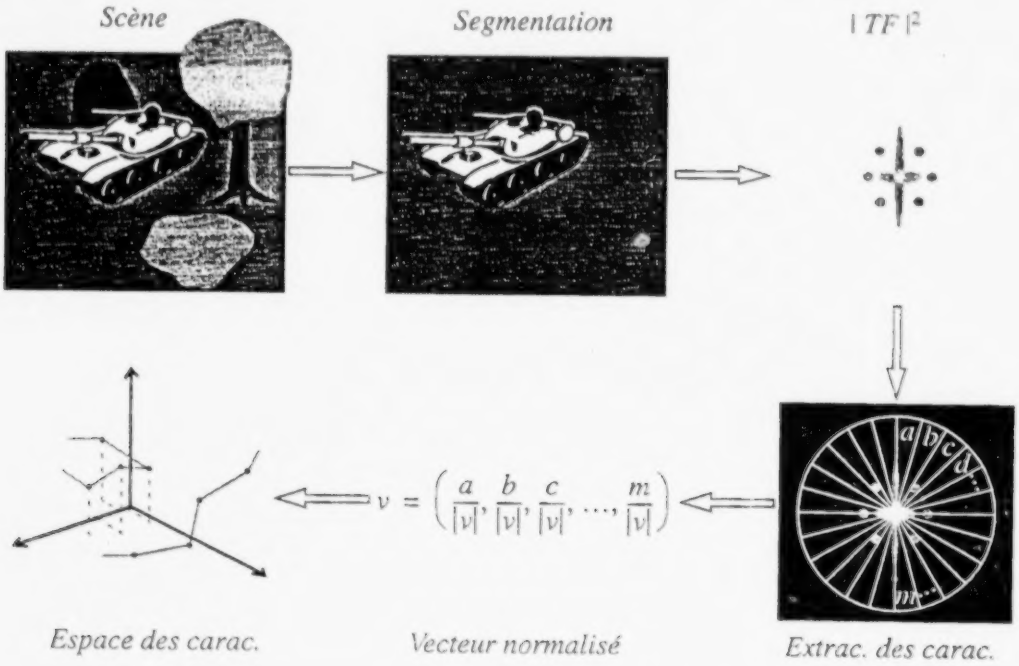


FIGURE 1.6 : Méthode des sections en pointe

Les simulations numériques ont été effectuées sur six types de véhicules. Ceux-ci ont été regroupés pour former deux classes. Des tests ont été effectués pour évaluer les performances de la méthode pour la rotation hors axe autour de l'axe vertical et pour les

changements d'échelle. Des problèmes concernant les changements d'échelle se sont manifestés, toutefois ceux-ci était dû à une erreur de calcul. Ne disposant pas d'images des objets à différentes tailles, les changements de taille ont été simulés. Toutefois le changement de la taille des objets par échantillonnage sur les images originales amène des problèmes de recouvrement spectral (ou aliasing). Le problème a été résolu en partie par le biais du filtrage moyen. Les simulations effectués sur 240 images (dont 120 faisaient partie de l'ensemble témoin) ont donné un taux d'erreur de classification inférieur à 5%.

1.4.1.2 Méthode des descripteurs de Fourier

La méthode des descripteurs de Fourier nécessite une bonne segmentation puisque les caractéristiques sont extraites sur le contour de l'objet. Pour assurer une bonne extraction des contours, une égalisation de l'histogramme a été effectuée sur chacune des images. De même un filtre médian de fenêtre 10 x 10 a été appliqué successivement à trois reprise sur chacune des images de façon à réduire le bruit.

L'extraction des contours a été réalisée à l'aide de la transformation en ondelette (TO). Le type d'ondelette utilisé est l'ondelette de spline quadratique parce que celle-ci représente une approximation de l'un des détecteurs optimaux de contours à savoir la première dérivée d'une gaussienne. Les composantes selon x et y de la TO sont utilisées pour définir un gradient. Le contour d'un objet est tracé à partir du point dans l'image correspondant à la plus forte variation d'intensité. Le tracé du contour se fait sur la base de l'information contenue dans le gradient. Étant donné le faible contraste des images, le tracé des contours n'a pu être fait de façon entièrement automatique et une intervention humaine a été necéssaire. La classification des objets se fait à partir de la représentation selon les descripteurs de Fourier dont la description est donné à la section 1.2.3.3. Le classificateur qui a été utilisé est du même type que celui de la méthode des sections en pointe.

Les expérimentations numériques ont été effectuées sur les mêmes véhicules que ceux utilisés pour les tests de la méthode des sections en pointe (SP), à l'exception du char AMX pour lequel seulement le quart des vues étaient disponibles. Tout comme pour la méthode SP, on considère des rotations hors plan autour de l'axe vertical seulement et les véhicules ont été regroupés selon deux classes. Sur 360 images, la méthode a permis de classer les véhicules avec moins de 10% d'erreur lorsque l'ensemble d'entrainement réuni toutes les vues à

intervale de 15° de rotation (120 images dans l'ensemble d'entrainement). Lorsque l'ensemble d'entrainement comprend toutes les vues à intervalle de 10°, l'erreur de classification est réduite alors à moins de 5%.

1.4.2 Travaux sur les CHC

La méthode des CHC est celle qui a été retenue en vue de la réalisation du modèle optique d'essai. Des tests préliminaires pour la mise en oeuvre optique de l'extraction des CHC ont été effectués au cours de la première partie du projet, toutefois beaucoup de travail restait à faire à ce niveau. Les résultats optiques préliminaires obtenus lors de la première partie du projet laissaient envisager certains problèmes au niveau de l'invariance sous rotation dans le plan. Au cours de la seconde partie du projet des efforts importants ont été consacrés en vue de faire l'extraction de signatures CHC valables. Ceci a amené des modifications sur la nature même des signatures CHC que l'on extrait désormais sur les objets eux même plutôt que sur leur TF. Suite à l'analyse des résultats préliminaires, les difficultés de la mise en oeuvre optique de l'extraction des CHC devenaient plus évidentes et ont été mieux cernées. Une simplification du problème global au niveau de la segmentation s'imposait pour permettre de consacrer d'avantage de temps à la réalisation des systèmes optiques. Les prochaines sous-sections décrivent le problème de RAC simplifié qui a été considéré pour la seconde partie du projet. On y fait également la description des deux montages optiques réalisés et on fait état des principaux résultats obtenus. Finalement un nouveau type de signature dérivé des CHC est introduit et des commentaires sont apportés à propos du modèle d'essai.

1.4.2.1 Nouvelle définition des CHC

Les premiers montages rudimentaires qui ont étés réalisés au début de la seconde partie du projet ont souligné la complexité de l'extraction optique des CHC. Une des constatations importantes apportées par les tests effectués à l'aide de ces montages concerne la transformée de Fourier optique. On se rappelle que pour obtenir des signatures CHC indépendantes de la position de l'objet, on avait convenu d'extraire les CHC sur la TF de l'objet. Pour obtenir une TF optique de qualité acceptable il faut utiliser de l'optique de qualité convenablement corrigée et assemblée avec précision. À défaut d'optique de précision, la TF optique se déforme et se décentre avec les déplacements de l'objet dans la scène et les changements d'orientation (rot. dans le plan). Or les CHC sont sensibles au

décentrage et la réalisation d'un système de TF optique de précision demandent beaucoup de temps. Pour ces raisons l'idée d'effectuer l'extraction des CHC sur les TF optiques des objets à été changé au profit de celle de l'extraction des CHC sur l'objet recentré. Cette stratégie permet de concentrer les efforts sur la vérification de la validité des CHC.

1.4.2.2 Problème considéré pour le modèle d'essai

Pour vérifier le comportement des CHC comme outil de reconnaissance, il convient de débuter le développement par un problème sur lequel on exerce un certain contrôle. Les tests sur des modèles réduits observés à l'aide d'une caméra vidéo ont donc été entrepris. Le problème des modèles réduits permet de simplifier la tâche de préparation des images et par le fait même permet de pouvoir consacrer d'avantage de ressources au problème de l'extraction optique des caractéristiques. Dans le cas des modèles réduits, l'arrière scène est un fond uniforme qui peut être éliminé par un simple seuil. Puisque l'extraction se fait sur les objets eux-mêmes plutôt que sur leur TF, il faut ajouter un étape de recentrage. Ceci se fait par la détermination des moments géométriques d'ordre un (centroïde) selon les lignes et les colonnes. Cette opération peut être affectée par la présence de bruit. Toutefois, un simple filtrage analogique suffit à éliminer les artefacts pouvant fausser l'évaluation du centroïde des objets.

Considérant les ressources et le temps disponibles et puisque que le mandat du projet est de se concentrer sur les aspects de reconnaissance, une approche systémique a été privilégiée afin de maîtriser l'extraction optique des caractéristiques. Le projet s'est donc orienté sur le problème illustré à la figure 1.7. Le type de scène considérée est constitué d'un modèle réduit monté sur un plateau rotatif et éclairé uniformément. Les images sont capturées avec une caméra vidéo, soit par le dessus dans le cas du problème 2D, soit par le côté pour le problème 3D. Le processeur Datacube, de concert avec un ordinateur SUN, effectue la numérisation, le filtrage analogique, le seuillage pour éliminer l'arrière scène et le recentrage de l'objet. Les images traitées par le Datacube doivent être ensuite présentées au système optique d'extraction des CHC. L'acquisition des résultats fournis par le système optique est effectuée par un ordinateur PC, lequel doit servir également de plate-forme pour le classificateur neuronal.

En principe, la représentation CHC est indépendante de l'orientation (rot. dans le plan) et de la taille (échelle) de l'objet et également de l'intensité de l'éclairement. Ainsi l'utilisation des CHC pour représenter des objets permet de contourner le problème des distorsions géométriques liées au problème 2D. Dans le cas du problème 3D, les changements d'aspect dues aux rotations hors plan doivent être pris en compte par le réseau de neurones. En pratique le réseau de neurones doit également fournir une certaine robustesse face aux variabilités dans l'évaluation des CHC. Heureusement la réalisation et l'entraînement du classificateur neuronal sont grandement facilités par l'importante compression de l'information opérée par les CHC.

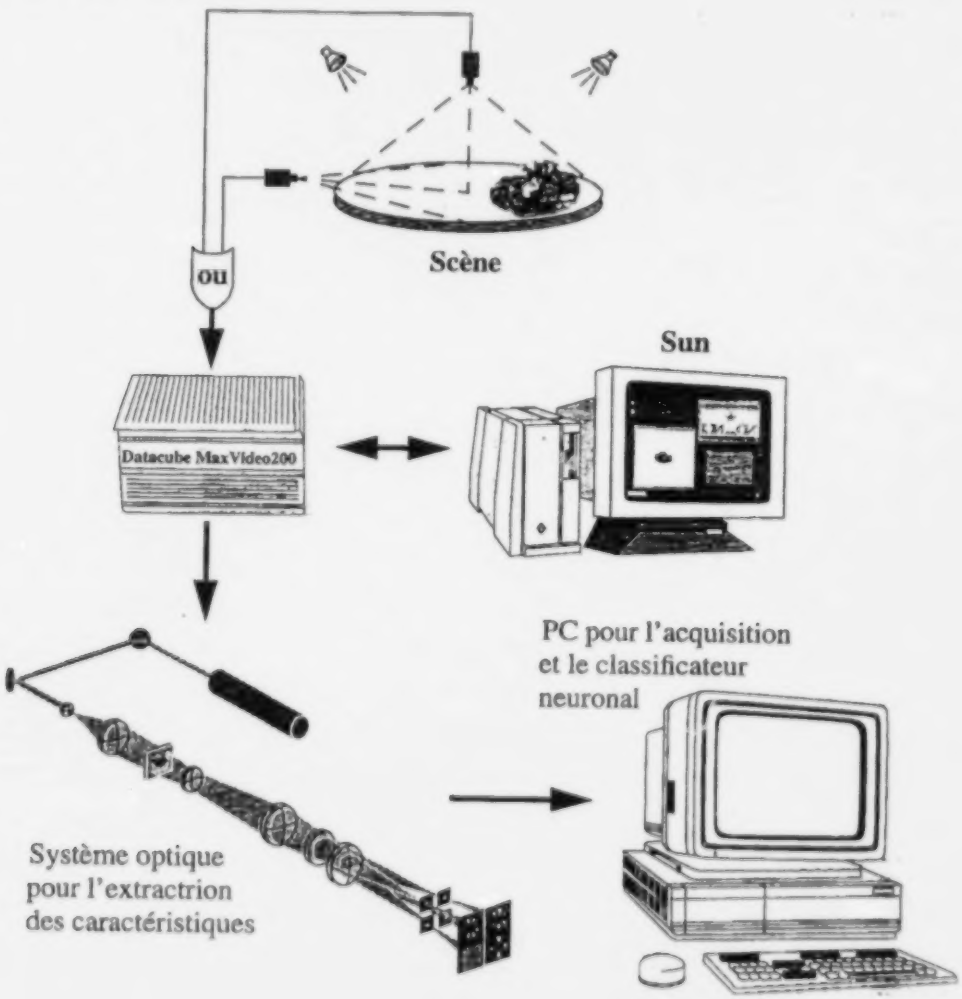


FIGURE 1.7 : Problème de RAC et modèle d'essai considéré.

1.4.2.3 Systèmes optiques pour l'extraction des CHC

À la fin de la première partie du projet, deux types de système optique avaient été proposés. Ceux-ci sont décrits à la section 1.3.2. Ces systèmes avaient été proposés pour faire l'extraction sur la TF des objets. Ceux-ci n'ont été réalisés que sous une forme rudimentaire et ont eu pour utilité de mettre en évidence les difficultés d'obtenir une TF optique de qualité avec un système optique non optimisé. Étant donné le changement introduit au niveau de la nature des CHC, de nouveaux systèmes optiques ont donc été considérés. Deux systèmes optiques ont été analysés, un premier système comporte deux ACL tandis que le second système est basé sur un réseau optique de diffraction.

1.4.2.3.1 Système à deux ACL

Ce système comporte deux ACL et est éclairé par une source incohérente. L'objet est affiché sur un premier ACL dont l'image est formé sur le second ACL avec un facteur de grandissement de 1X, à l'aide d'un objectif très bien corrigé pour les aberrations (notamment pour la distorsion). La lumière transmise par les deux ACL est recueillie sur un photodétecteur, lequel est équipé d'un circuit d'amplification. L'amplification des signaux électriques provenant des détecteurs se fait par des circuits à deux étages. Malgré le peu de temps consacré à l'ajustement de ce montage, des résultats intéressants ont été obtenus. C'est sur ce système que nos meilleures signatures CHC optiques ont été obtenues. Ce système est illustré à la figure 1.8.

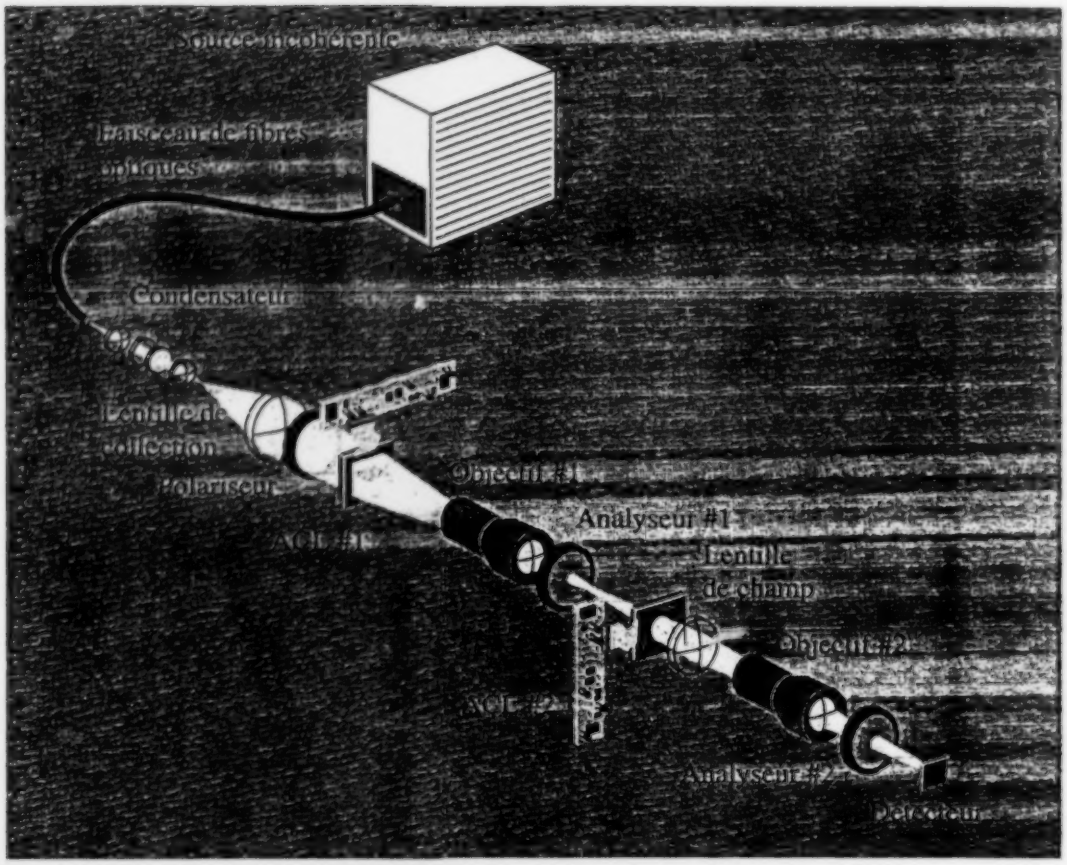


FIGURE 1.8 : Système à deux ACL

1.4.2.3.2 Système à réseau de diffraction

Le système réalisé est illustré à la figure 1.9. Celui-ci comprend un ACL éclairé par un faisceau laser de même qu'un système d'imagerie qui permet de produire l'image du ACL dans le plan contenant les rosettes. Le réseau de diffraction est placé dans le parcours optique juste en avant de la dernière lentille et permet ainsi de produire des répliques de l'image affichée par l'ACL. Les rosettes sont placées sur une surface verticale et peuvent être ajustées individuellement à l'aide d'un dispositif amovible. Le plan des filtres spatiaux, destinés à éliminer les effets de l'échantillonnage spatial de l'ACL, est placé un peu en amont du plan des rosettes. Les signaux optiques transmis par les rosettes sont recueillis à l'aide de

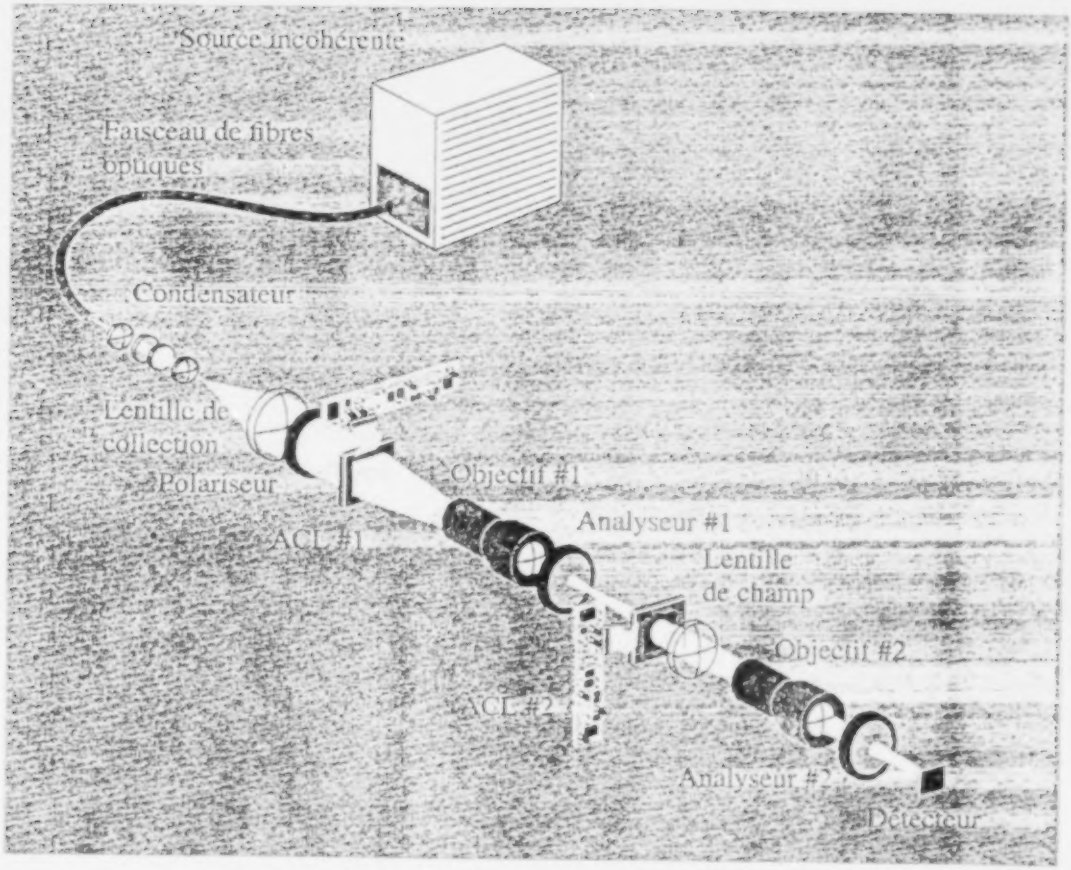


FIGURE 1.8 : Système à deux ACL

1.4.2.3.2 Système à réseau de diffraction

Le système réalisé est illustré à la figure 1.9. Celui-ci comprend un ACL éclairé par un faisceau laser de même qu'un système d'imagerie qui permet de produire l'image du ACL dans le plan contenant les rosettes. Le réseau de diffraction est placé dans le parcours optique juste en avant de la dernière lentille et permet ainsi de produire des répliques de l'image affichée par l'ACL. Les rosettes sont placées sur une surface verticale et peuvent être ajustées individuellement à l'aide d'un dispositif amovible. Le plan des filtres spatiaux, destinés à diminuer les effets de l'échantillonnage spatial de l'ACL, est placé un peu en amont du plan des rosettes. Les signaux optiques transmis par les rosettes sont recueillis à l'aide de

photodétecteurs fixés juste en arrière des rosettes. L'amplification des signaux électriques provenant des détecteurs se fait par des circuits à deux étages. Les circuits de préamplification sont placés près de leur détecteur respectif tandis que les circuits d'amplification sont réunis sur une même plaquette de circuit imprimée.

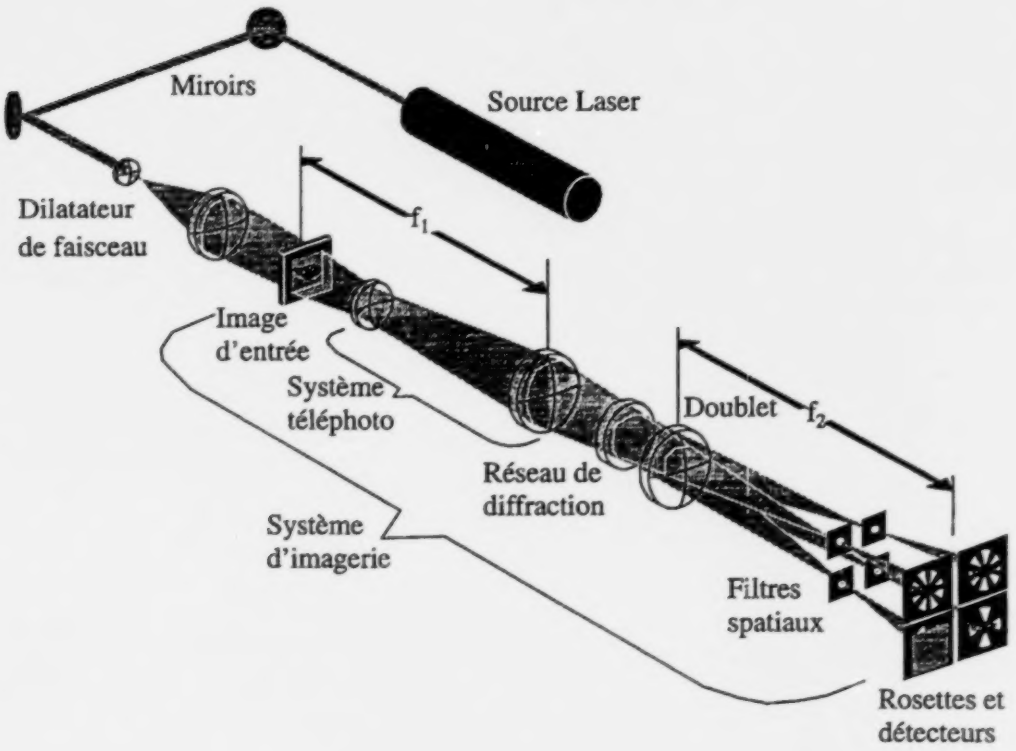


FIGURE 1.9 : Système optique d'extraction des CHC avec réseau de diffraction.

La réalisation de ce système optique a demandé beaucoup d'effort et de temps en raison de la précision requise pour l'ajustement des composants. Il s'agit d'un système à éclairement cohérent et les images produites dans le plan des rosettes sont entachées de bruit. Les rosettes ont été réalisées grâce à la technologie de l'écriture par faisceau d'électron sur masque en chrome, ce qui a impliqué un temps d'attente important. Comble de malheur, celles-ci se sont ensuite révélées inutilisable en raison d'une erreur qui s'est glissée lors de la

création des fichiers. Le résultat de cette erreur est l'enregistrement de rosettes en $(1+cos)^{1/2}$ plutôt qu'en $1+cos$, ce qui est plus que suffisant pour anéantir les résultats. Le système optique de l'extraction des CHC basé sur un réseau de diffraction n'a donc pas donné les résultats attendus. Ceci a eu une influence importante sur l'évolution du projet puisqu'une grande partie des ressources ont été affectées à ce montage.

Malgré l'erreur sur l'enregistrement des rosettes qui a elle seule suffit à rendre inopérable le système à réseau de diffraction, la piètre qualité des images obtenues laisse croire que de toute façon il serait difficile d'obtenir des résultats de qualité avec un tel système. Il faudrait entre autres faire une conception optique plus élaborée et assembler le système dans un environnement exempt de poussière.

Malgré tout, l'expertise développée sur la réalisation des rosettes sur masque en chrome permettrait d'améliorer et de simplifier le montage à deux ACL présenté plus haut, la technologie utilisée pour la réalisation des rosettes ayant démontrée dans l'ensemble de bons résultats.

1.4.2.3.3 Comparaison des deux systèmes optiques

Le système à deux ACL comporte de nombreux avantages relativement à celui basé sur le réseau de diffraction. Tout d'abord le système à deux ACL fonctionne avec une illumination incohérente. Ceci est un réel avantage puisque l'illumination cohérente du système à réseau de diffraction fait en sorte que les répliques de l'image d'entrée sont bruitées. La qualité des images produites par le système à deux ACL est bien supérieure à celle du système à réseau de diffraction.

Pour le système à deux ACL la compensation des non uniformités de l'éclairement qui sont dues à la source et au vignettage se fait facilement par superposition d'une figure de compensation sur les rosettes affichées par le second ACL. Dans le cas du système à réseau de diffraction la compensation des non-uniformités demanderait l'ajout de masques supplémentaires.

Le système à deux ACL est de taille beaucoup plus petite que celle du système à réseau de diffraction. Des réductions des dimensions seraient possible pour les deux systèmes, toutefois le système à deux ACL serait celui des deux qui semble le plus facilement miniaturisable. La fonctionnalité du système à deux ACL pourrait probablement être reproduite dans un système de taille guère plus importante que celle d'un ACL (voir la section 3.3.2).

1.4.2.4 CHC angulaires

Les CHC angulaires constituent de nouvelles caractéristiques qui sont dérivées des CHC. Tel qu'illustré à la figure 1.10, les valeurs associées aux rosettes d'un ordre harmonique donné varient sinusoïdale avec l'angle de rotation (dans le plan) des objets. Le déphasage relatif des courbes sinusoïdales est une propriété propre à l'objet. L'ensemble des angles de déphasage convenablement référencé (voir la section 2.1.4) constitue une signature invariante sous rotation dans le plan et sous échelle. Celle-ci a été appelé *signature CHC angulaire* (CHCA). En pratique les signatures CHCA sont constituées d'un nombre restreint de CHCA. Les signatures CHCA sont calculées avec les valeurs des paires de rosettes CHC et ne requiert donc pas de rosettes ou fonctions supplémentaires. En se basant sur les résultats obtenus, il semble que les signatures CHCA présenteraient des différences plus marquées pour des objets distincts comparativement à leurs homologues CHC.

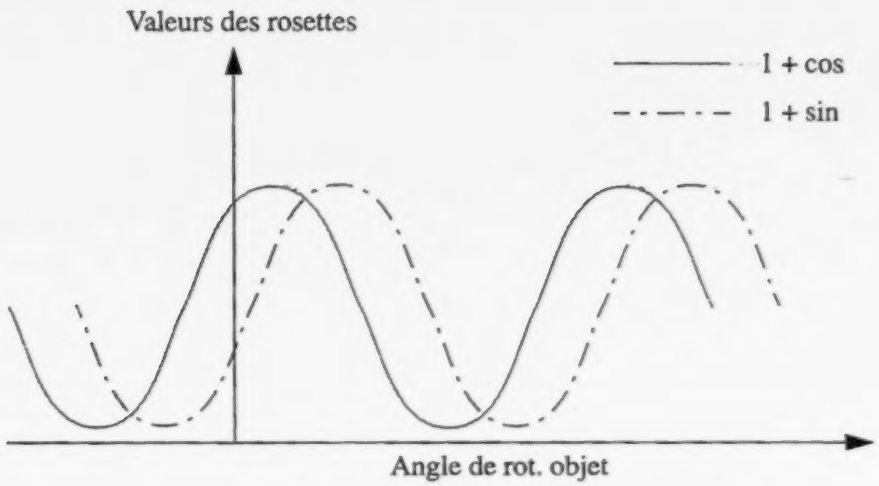


FIGURE 1.10 : Courbes types des valeurs des rosettes en fonction de l'angle de rotation de l'objet.

1.4.2.5 Résultats avec les CHC et les CHCA

Les tests ont été effectués à l'aide de véhicules modèles dont la taille varie de 45 à 180 pixels selon l'échelle. Deux séries de tests ont été réalisées, une première numérique et une seconde optique. Les tests numériques visaient à confirmer la capacité de discrimination des CHC évalués sur les objets eux-même plutôt que sur leur TF (contrairement aux résultats obtenus avec les CHC lors de la première partie du projet) et à évaluer les CHCA comme nouvel outil de reconnaissance. Les tests optiques ont été réalisés à l'aide du montage à deux ACL présenté à la figure 1.8.

Les tests numériques portant sur la reconnaissance 2D ont permis de démontrer l'invariance des signatures en rotation dans le plan, intensité et échelle sur les images test, à la fois pour les CHC et les CHCA. Les CHCA ont requis un nombre plus restreint d'images d'entraînement et d'itérations pour obtenir la reconnaissance complète des véhicules modèles relativement à ce qui a été nécéssaire pour les CHC. Pour les CHCA, l'ensemble d'entraînement comportait une seule signature par objet alors qu'il en comportait 2 par objet dans le cas des CHC. L'entraînement a nécessité 260 itérations dans le cas des CHCA et de l'ordre de 7200 pour les CHC. Les tests ont porté sur 2072 versions de huit objets tournés sur

360 degrés dans le plan et montrant des facteurs d'échelle variant entre 0.5 et 2.0. Toutes les cibles ont été identifiées correctement tant avec les CHCA qu'avec les CHC.

Les résultats numériques sur le problème 3D ont, eux aussi, permis la reconnaissance complète des objets. Ici encore, les CHCA ont requis un nombre restreint d'itérations (185 comparativement à environ 10000 pour les CHC). De plus, les CHCA n'ont nécessité l'utilisation que d'un seul réseau de neurones. Les CHC ont quant à eux nécessité l'utilisation de trois réseaux de neurones organisés en une structure décisionnelle hiérarchique.

Pour le problème 3D, les CHCA ont permis une reconnaissance complète de tous les objets tandis que les CHC ont montré un taux de reconnaissance élevé (97%). Les tests ont été effectués sur un total de 96 images montrant différentes vues de quatre objets (modèles réduits de véhicules). Le nombre d'images de l'ensemble d'entraînement est d'environ 20 pour les CHC et 47 pour les CHCA. Il est à noter que les réseaux de neurones utilisés pour la reconnaissance sont des perceptrons à une seule couche cachée et comportent un nombre restreint de neurones (<28) dans la couche cachée.

Les expériences optiques réalisées pour le problème 2D ont permis d'obtenir des signatures globalement stables pour les six premiers ordres de CHCA, à la fois en rotation dans le plan et sous changement d'échelle. Les données transmises au réseau de neurones ont permis la reconnaissance de toutes les images (96) à l'aide de 6 images d'entraînement dans un nombre restreint d'itérations (62). De plus, une indication stable de la rotation de l'objet a pu être obtenue à l'aide de l'information obtenue du premier ordre.

Les CHC ont montré une signature légèrement moins stable mais toutefois utilisable pour les quatre premiers ordres. Le taux de reconnaissance s'est avéré élevé (99%) en 2380 itérations pour 11 images d'entraînement. Les tests optiques sur la reconnaissance 3D n'ont pas été effectué faute de temps.

1.4.2.6 Modèle d'essai

Le modèle d'essai n'a pu être complété par manque de temps. Le système à réseau de diffraction autour duquel devait s'orchestrer les autres systèmes n'a pu être mis en opération.

Toutefois, le système à deux ACL a permis d'obtenir des résultats qui semblent prometteurs particulièrement en ce qui concerne les CHCA. Ce dernier système a été réalisé à la toute fin du projet alors qu'il ne restait que très peu de temps disponible. Il n'a donc pas été possible d'optimiser le montage et de procéder à l'adaptation des logiciels qui ont été développés dans la perspective du système à réseau de diffraction. Le système à réseau de diffraction devait faire le calcul de toutes les valeurs des rosettes simultanément et évidement les logiciels pour la préparation des images, pour l'acquisition et le traitement des résultats optiques avaient été construit en conséquence. Dans sa configuration actuelle, le système à deux ACL fait le calcul d'une seule valeur de rosette par cycle vidéo. Pour pouvoir compléter le modèle d'essai il faudrait adapter tout les logiciels déjà réalisés de manière à tenir compte du fait que la signature est générée sur plusieurs cycles vidéo plutôt que sur un seul.

Une bonne partie du modèle d'essai à tout de même été complétée. Les réseaux de neurones ont été mis en oeuvre sur plate-forme PC de même que le logiciel d'acquisition des résultats optiques et son interface. Le système simplifié de pré-traitement des images brutes sur plate-forme Datacube a également été complété.

1.5 Conclusion

Bien que le modèle d'essai prévu dans ce projet n'a pu être complété en raison de manque de temps et de ressources, le projet a permis le développement et l'étude de trois méthodes d'extraction de caractéristiques de cibles et de la méthode de corrélation DCCF. Les méthodes des CHC, des sections en pointe et des descripteurs de Fourier ont fait l'objet d'études approfondies par le biais de simulations numériques. Les simulations ont montré que dans un contexte simple, la reconnaissance d'objet 3D à partir d'information 2D est possible.

La méthode des CHC est celle qui a été retenue pour la réalisation optique du modèle d'essai. L'architecture optique de l'extraction des CHC utilisant deux afficheurs s'est avérée la meilleure alternative. La reconnaissance optique complète en 2D à l'aide de réseaux de neurones de type perceptron à trois couches a pu être accomplie pour un problème de RAC simple. Les expériences optiques sur le problème 3D n'ont pu être entamées par manque de temps. Les données optiques déjà obtenues en RAC 2D pourraient être améliorées par différentes stratégies. Les différents aspects étudiés lors de ce projet ont permis d'identifier

ces stratégies. La méthode DCCF a été étudiée seulement dans la première partie du projet, les travaux ont été interrompus en raison des résultats mitigés obtenus.

Le projet a également impliqué l'étude d'algorithmes de préparation des images précédant l'extraction des caractéristiques. Plusieurs algorithmes simples ont été mis en oeuvre sur le processeur Datacube. D'autres, comme la transformation en ondelettes, ont fait l'objet de simulations. La classification de signatures soit à l'aide de RNA d'architectures simples, soit par la méthode des trajectoires dans l'espace des caractéristiques, ont également fait l'objet d'études.

Les travaux effectués, ont démontré que certaines des méthodes à extraction des caractéristiques étudiées au cours du projet présentent de l'intérêt pour la reconnaissance des formes. La segmentation associée à la reconnaissance d'un objet requiert un effort additionnel. Finalement, il nous apparaît clairement que la solution du problème global de RAC, d'intérêt pour le CRDV, implique des ressources beaucoup plus importantes que celles qui ont été consacrées au projet jusqu'ici.

ERRATA

La définition des descripteurs de Fourier présentée dans ce rapport diffère de celle présentée dans le rapport de la première partie du projet.

Dans le premier rapport il est mentionné que la mise en oeuvre du filtre pseudo-médian avait été réalisée. Il s'agit d'une erreur, le filtre pseudo-médian n'ayant pas été mis en oeuvre sur Datacube.

Première partie: Contribution de l'INO

CHAPITRE 2

Signatures CHC

Dans ce chapitre la nouvelle définition des coefficients d'harmoniques circulaires (CHC) est présentée. Cette nouvelle définition tiens compte du fait que l'idée de faire l'extraction des CHC sur les transformées de Fourier (TF) a été abandonnée. L'extraction des CHC se fait plutôt sur l'image de l'objet centré. Le chapitre comprend également une évaluation théorique des erreurs d'estimation des CHC obtenus optiquement, ainsi qu'une généralisation de la formulation des CHC utilisés pour la mise en oeuvre optique. Un nouveau type de caractéristique dérivé des CHC sera également présenté. Finalement les algorithmes développés pour compenser en partie les erreurs liées à la mise en oeuvre optique sont présentés.

2.1 Signatures CHC

L'estimation des CHC de façon optique est une opération délicate. Pour obtenir les meilleurs résultats possibles il convient de bien maitriser les éléments de théorie se reportant aux CHC. Le développement de la théorie liés aux CHC est fait dans les sous-sections suivantes. Le cheminement conduit à des relations qui seront utilisées pour la correction de certaines erreurs associées aux mesures optiques. De plus la démarche à permis de trouver un nouveau type de caractéristiques dérivées des CHC.

2.1.1 Nouvelle définition des CHC

Dans le rapport de mi-projet, une définition des coefficients d'harmonique circulaire (CHC) avait été donnée. De manière à tenir compte des nouvelles réalités, un élargissement de la définition des CHC s'impose. Auparavant le terme CHC référait au module carré du descripteur de Fourier-Mellin du module carré de la TF. On faisait intervenir la TF dans le but d'obtenir l'invariance sous translation. Toutefois la TF n'est pas indispensable si l'on dispose

d'une méthode fiable qui permette d'identifier correctement un point unique de l'objet quelque soit sa taille et son orientation (rotation dans le plan). Dans les situations simples où les limites de l'objet sont bien définies, les moments d'ordre un selon deux directions orthogonales permettent d'identifier le centre géometrique (centroïde) de l'objet avec une bonne fiabilité. Puisque la mise en oeuvre optique de l'extraction des CHC sur les TF se heurte à plusieurs difficultés, il a été jugé moins risqué de faire l'extraction sur les objets et non sur leur TF. On va donc redéfinir le coefficient d'harmonique circulaire d'ordre n,m d'un objet $f(r,\phi)$ comme étant le module carré du descripteur de Fourier-Mellin [1] d'ordre $n-1$, m de sa version $f_c(r,\phi)$ centrée sur son centre géométrique:

$$H_{nm}(f) \equiv \left| \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} f_c(r,\theta)\, e^{-im\theta} r^{n+1} d\theta dr \right|^2 . \tag{1}$$

Ainsi défini $H_{nm}(f)$ est un scalaire. En utilisant la décomposition de $f_c(r,\theta)$ en harmoniques circulaires :

$$f_c(r,\theta) = \sum_{m=-\infty}^{\infty} f_{cm}(r)\, e^{im\theta} \tag{2}$$

$$\text{où } f_{cm}(r) = \int_0^{2\pi} f_c(r,\theta)\, e^{-im\theta} d\theta , \tag{3}$$

on démontre sans trop de difficulté que $H_{nm}(f)$ ne dépend pas de l'orientation de l'objet dans le plan [1,6]. De la même façon, on démontre que les CHC normalisés définis comme étant :

$$\bar{H}_{nm}(f) \equiv \frac{H_{nm}(f)}{H_{n0}(f)} \leq 1 , \tag{4}$$

ne dépendent pas de la taille de l'objet [1,6].

2.1.2 Formulation pour la mise en oeuvre optique

La nouvelle définition des CHC fait intervenir une somme intégrale impliquant des arguments complexes. En principe l'intégrale complexe pourrait être évaluée optiquement puisqu'elle se ramène à la valeur à l'origine d'une transformée de Fourier [7,8].

$$H_{nm}(f) = \left| \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} f_c(r,\theta)\, e^{-im\theta} r^{n+1} dr d\theta \right|^2$$

$$= \left| \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} \left(f_c(r,\theta)\, e^{-im\theta} r^n \right) e^{i2\pi r\rho\cos(\theta-\phi)} r dr d\theta \Big|_{\rho=0} \right|^2 \qquad (5)$$

$$= \left| \frac{1}{\pi} TF\left(f_c(r,\theta)\, e^{-im\theta} r^n \right) \Big|_{\rho=0} \right|^2 .$$

Toutefois l'évaluation optique de la transformée de Fourier est difficile pour deux raisons. Premièrement on doit disposer d'un modulateur spatial de lumière modulable à la fois en phase et en amplitude pour pouvoir afficher la fonction complexe $r^n f_c(r,\theta)\, e^{-im\theta}$. Deuxièmement, le bruit inhérent au système optique cohérent risque de détériorer grandement les résultats puisque l'information pertinente occupe une faible portion (Point central) du plan de la TF. La difficulté à évaluer les CHC par TF optique a d'ailleurs été vérifiée au cours de la première partie du projet. Un système avait été assemblé et malgré tous les efforts aucun résultat acceptable n'a été obtenu.

Heureusement, il y a une autre méthode pour évaluer optiquement les CHC puisque l'on peut exprimer les H_{nm} d'une façon qui se prête mieux à la mise en oeuvre optique. En effet on peut démontrer [6] que :

$$H_{nm}(f) = 4\left[C_{nm} - \frac{1}{2} C_{n0} \right]^2 + 4\left[S_{nm} - \frac{1}{2} C_{n0} \right]^2 \qquad (6)$$

avec $$C_{nm}(f) = \frac{1}{\pi} \int_0^{\infty} \int_0^{2\pi} f_c(r,\theta) \left(\frac{1+\cos(m\theta)}{2} \right) r^{n+1} d\theta dr$$

$$S_{nm}(f) = \frac{1}{\pi}\int_0^{\infty}\int_0^{2\pi} f_c(r,\theta)\left(\frac{1+sin(m\theta)}{2}\right)r^{n+1}d\theta dr.$$

Les intégrales impliquées dans cette formulation font intervenir uniquement des fonctions réelles positives ($f(r,\theta)$ = image $\Rightarrow$ $\mathbb{R}^+$). Selon les relations 5 et 6, $H_{n0} = C_{n0}^{\ 2}$. Aussi les CHC normalisés sont donnés par la relation suivante :

$$H_{nm} = 4\left[\frac{C_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right]^2 + 4\left[\frac{S_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right]^2. \qquad (7)$$

L'évaluation des C'_{nm} et S'_{nm} peut se faire optiquement par projection de l'objet $f_c(r,\theta)$ sur des masques en [1+cos(mθ)]/2 et [1+sin(mθ)]/2. Cette formulation pour les CHC est à la base des travaux présentés dans ce rapport. Une formulation plus générale a également été développée pour permette la correction de certaines erreurs expérimentales. Avant de présenter cette version plus générale il faut développer d'avantage les expressions des C'_{nm} et S'_{nm}.

2.1.3 Effet de la rotation de l'objet sur les valeurs C'_{nm} et S'_{nm}

Considérons un objet $f_c(r,\theta)$ et l'une de ses versions, $f_c(r,\theta+\phi)$, tournée relativement à son centre géométrique. Pour cette version tournée

$$C_{nm}(f;\phi) = \frac{1}{\pi}\int_0^{\infty}\int_0^{2\pi} f_c(r,\theta+\phi)\left(\frac{1+\cos(m\theta)}{2}\right)r^{n+1}d\theta dr. \qquad (8)$$

La décomposition de l'objet selon ses composantes d'harmoniques circulaires est donnée par la relation suivante :

$$f_c(r,\gamma) = \sum_{k=-\infty}^{\infty} f_{ck}(r)\, e^{ik\gamma} \tag{9}$$

$$\text{avec} \quad f_{ck}(r) \equiv \frac{1}{\pi}\int_0^{2\pi} f_c(r,\gamma)\, e^{-ik\gamma} d\gamma\,.$$

En remplaçant cette dernière relation dans l'argument de l'intégrale de $C'_{nm}(f;\phi)$ on obtient

$$C_{nm}(f;\phi) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} e^{ik\phi} F_{nk}\,[I_{0k} + I_{mk}] \tag{10}$$

$$\text{où} \qquad F_{nk} \equiv \int_0^{\infty} f_{ck}(r)\, r^{n+1} dr \tag{11}$$

$$I_{mk} \equiv \int_0^{2\pi} \cos(m\theta)\, e^{ik\theta} d\theta\,. \tag{12}$$

On démontre aisément que $I_{0k} = 0$ et

$$I_{mk} = \pi\,(2\delta_{0,k} + \delta_{k,m} + \delta_{k,-m})\,, \tag{13}$$

où $\delta_{i,j}$ est le delta Kronecker qui vaut 1 lorsque i = j et 0 dans le cas contraire. Cette dernière relation ramène la somme de la relation (10) en une somme de trois termes, à savoir :

$$C_{nm}(f;\phi) = F_{n0} + \frac{1}{2}F_{nm}\, e^{im\phi} + \frac{1}{2}F_{n-m}\, e^{-im\phi}\,. \tag{14}$$

Du fait que les C'_{nm} sont à valeur réelle, on déduit que :

$$F_{n-m} = F^*_{nm} \qquad \forall n, m \in entier. \tag{15}$$

Par conséquent :

$$C_{nm}(f;\phi) = F_{n0} + \Re\left(F_{nm}e^{im\phi}\right) = F_{n0} + |F_{nm}|\cos(m\phi + \phi_{nm}) . \tag{16}$$

Où $\Re(z)$ désigne la partie réelle de la variable complexe z. Selon la relation (16) :

$$F_{n0} = \frac{1}{2}\cdot C_{n0}(f,\phi) = \frac{1}{2}\cdot C_{n0}(f) . \tag{17}$$

Cette dernière relation conduit au résultat final suivant :

$$C_{nm}(f;\phi) - \frac{1}{2}\cdot C_{n0}(f,\phi) = |F_{nm}|\cos(m\phi + \phi_{nm}) . \tag{18}$$

En procédant de façon similaire on déduit

$$S_{nm}(f;\phi) - \frac{1}{2}\cdot C_{n0}(f,\phi) = |F_{nm}|\sin(m\phi + \phi_{nm}) . \tag{19}$$

On voit que les paramètres C'_{nm} et S'_{nm} varient sinusoïdalement avec la rotation de l'objet dans le plan. Les expressions (18) et (19) vont être d'une grande utilité pour l'élaboration d'une généralisation de la formulation des CHC.

2.1.4 Introduction des CHC angulaires

Jusqu'à présent, les signatures CHC concernaient l'information sur l'amplitude des CHC. C'est-à-dire que le CHC est un coefficient représentatif de la proportion d'une harmonique circulaire contenue dans un objet. Ainsi, la signature CHC représente les pondérations relatives de la présence des différents ordres d'harmoniques circulaires dans un objet.

Les équations (18) et (19) font ressortir un aspect important du calcul des CHC. Les CHC sont obtenus par une combinaison d'intégrales avec des fonctions primitives en sinus et en cosinus. Lors du calcul des CHC, les termes en sinus et cosinus sont couplés afin de faire

disparaître la dépendance angulaire des intégrales C'_{nm} et S'_{nm} pour ne laisser que l'information sur le module $|F^{nm}|$. L'information sur la phase des paramètres F_{nm} peut toutefois être utilisée pour extraire une information supplémentaire. On veut donc évaluer la valeur ϕ_{nm} des phases propres à l'objet qui apparaissent dans les relations (18) et (19). Selon ces relations

$$A_{nm} \equiv \operatorname{atan}\left(\frac{S_{nm} - \frac{1}{2}\cdot C_{n0}}{C_{nm} - \frac{1}{2}\cdot C_{n0}}\right) = \operatorname{atan}\left(\frac{|F_{nm}|\sin(m\phi + \phi_{nm})}{|F_{nm}|\cos(m\phi + \phi_{nm})}\right) = m\phi + \phi_{nm} \qquad (20)$$

puisque la valeur de l'angle de rotation ϕ n'est pas connue à priori, il faut l'éliminer. Or selon la relation (20) :

$$A^{m_0}{}_{nm} \equiv A_{nm} - \frac{m}{m_0}A_{nm_0} = [m\phi + \phi_{nm}] - \frac{m}{m_0}[m_0\phi + \phi_{nm_0}] = \phi_{nm} - \frac{m}{m_0}\phi_{nm_0} \qquad (21)$$

Cette dernière relation est totalement indépendante de l'orientation de l'objet et elle définie ce qu'on appelle le ***coefficient d'harmonique circulaire angulaire*** (CHCA) d'ordre n m. Contrairement au CHC, dont les valeurs diminuent avec l'ordre, les CHCA montrent des valeurs variant entre 0 et 360 degrés quelque soit l'ordre. Une plus grande plage de variation possible est donc anticipée de la part des CHCA. Il est à noter finalement que l'information des CHC et des CHCA bien qu'évaluée à partir des mêmes paramètres, est différente. Ainsi ces informations sont complémentaires, et l'utilisation des deux informations couplées devrait s'avérer un outil supplémentaire pour obtenir de meilleures performances en reconnaissance.

2.2 Incertitudes liées à l'évaluation optique des CHC

Selon la relation (7), pour évaluer optiquement les valeurs des CHC normalisés on a besoin de trois mesures. Des incertitudes sur l'une ou l'autre de ces mesures entraînent des erreurs sur la valeur des CHC. L'incertitude des mesures peut être causée par le bruit ou encore due à des

imprécisions dans le positionnement des composants. Il est utile de faire une évaluation théorique des incertitudes susceptibles de ce manifester. Cela permet d'identifier les points faibles de la méthode et de prévoir des mécanismes de compensation. Dans les prochaines sous-sections, une évaluation théorique de la précision de la mise en oeuvre optique sera présentée.

2.2.1 Évaluation de la borne supérieure de l'incertitude relative

On utilise la différentielle totale pour estimer l'incertitude sur l'évaluation des CHC. Partant de la relation 7, on obtient :

$$dH_{nm} = \frac{\partial H_{nm}}{\partial C_{nm}} dC_{nm} + \frac{\partial H_{nm}}{\partial S_{nm}} dS_{nm} + \frac{\partial H_{nm}}{\partial C_{n0}} dC_{n0} \tag{22}$$

$$= \frac{8}{C_{n0}^2}\left[\left(C_{nm} - \frac{1}{2}C_{n0}\right)dC_{nm} + \left(S_{nm} - \frac{1}{2}C_{n0}\right)dS_{nm}\right]$$

$$-\frac{4}{C_{n0}^2}\left[C_{nm} + S_{nm} - C_{n0} + \frac{1}{2}H_{nm}C_{n0}\right]dC_{n0} .$$

On peut fixer une limite supérieure pour l'incertitude relative, en utilisant l'inégalité triangulaire [9]:

$$|dH_{nm}| \le \left|\frac{\partial H_{nm}}{\partial C_{nm}}\right||dC_{nm}| + \left|\frac{\partial H_{nm}}{\partial S_{nm}}\right||dS_{nm}| + \left|\frac{\partial H_{nm}}{\partial C_{n0}}\right||dC_{n0}| . \tag{23}$$

En considérant que les incertitudes dC'_{nm}, dS'_{nm} et dC'_{n0} sont de même importance, alors la relation (23) se résume à l'expression suivante :

$$|dH_{nm}| \le 4\left|\frac{dC_{n0}}{C_{n0}}\right|\left[3\left|\frac{C_{nm} - \frac{1}{2}C_{n0}}{C_{n0}}\right| + 3\left|\frac{S_{nm} - \frac{1}{2}C_{n0}}{C_{n0}}\right| + \frac{1}{2}H_{nm}\right] . \tag{24}$$

Selon la relation (7) :

$$4\left[\frac{C_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right]^2 \le 4\left[\frac{C_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right]^2 + 4\left[\frac{S_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right]^2 = \bar{H}_{nm} \le 1,$$

aussi :

$$\left|\frac{C_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right| \le \frac{\sqrt{\bar{H}_{nm}}}{2}, \tag{25}$$

et de même, on a que :

$$\left|\frac{S_{nm}-\frac{1}{2}C_{n0}}{C_{n0}}\right| \le \frac{\sqrt{\bar{H}_{nm}}}{2}. \tag{26}$$

En utilisant ces deux derniers résultats dans la relation (24) on obtient finalement la relation suivante pour la borne supérieure de l'incertitude relative :

$$\left|\frac{d\bar{H}_{nm}}{\bar{H}_{nm}}\right| \le \left[\frac{12}{\sqrt{\bar{H}_{nm}}}+2\right]\left|\frac{dC_{n0}}{C_{n0}}\right|. \tag{27}$$

Cette dernière relation, bien qu'approximative, donne une idée des résultats auxquels on peut s'attendre. La relation (27) comprend un terme inversement proportionnel à la racine de la valeur du CHC normalisé. Ceci fait en sorte que l'on doit s'attendre à observer des incertitudes relatives importantes pour les CHC d'ordres harmoniques élevées (m grand). En effet la valeur des CHC décroit asymptotiquement avec l'ordre harmonique m. Aussi pour des ordres harmoniques élevées pour lesquelles $\bar{H}_{nm}$ est de l'ordre 10^{-2}, l'incertitude relative sur l'estimation de $\bar{H}_{nm}$ peut être jusqu'à 100 fois supérieure à l'incertitude relative sur la mesure des C_{nm} et S_{nm}. Même pour les ordres harmoniques faibles, on doit s'attendre à des incertitudes relativement importantes sur l'évaluation des CHC, puisque $\bar{H}_{nm} < 1$. Pour un cas idéal où

$H_{nm} \sim 1$, la borne supérieure de l'incertitude relative correspond à environ 14 fois l'incertitude relative sur la mesure de C'_{n0}. L'évaluation optique des CHC necéssite donc des mesures de bonne précision et seuls les ordres harmoniques de faibles valeurs ($m < \sim 10$) sont accesibles.

En procédant de façon similaire il est possible de trouver également une expression pour la borne supérieure de l'erreur d'évaluation des CHCA. Il faut d'abord déterminer l'expression de la différentielle totale

$$dA^{m_0}{}_{nm} \equiv dA_{nm} - \frac{m}{m_0} dA_{nm_0} \leq \left|dA_{nm}\right| + \frac{m}{m_0}\left|dA_{nm_0}\right| \tag{28}$$

Pour compléter la relation il faut une expression pour la borne supérieure des dA_{nm} en fonction des mesurables C'_{nm}, S'_{nm} et C'_{n0}. Or

$$dA_{nm} = \frac{1}{1+u^2}\left(\frac{\partial u}{\partial S_{nm}} dS_{nm} + \frac{\partial u}{\partial C_{nm}} dC_{nm} + \frac{\partial u}{\partial C_{n0}} dC_{n0}\right) \tag{29}$$

$$\text{où } u \equiv \frac{S_{nm} - \frac{1}{2}\cdot C_{n0}}{C_{nm} - \frac{1}{2}\cdot C_{n0}}$$

En considérant comme précédemment que les incertitudes dC'_{nm}, dS'_{nm} et dC'_{n0} sont de même importance, alors la borne supérieure de l'erreur peut être exprimée sous la forme suivante :

$$\begin{aligned} \left|dA_{nm}\right| &\leq \frac{\left|dC_{n0}\right|}{1+u^2}\left(\left|\frac{\partial u}{\partial S_{nm}}\right| + \left|\frac{\partial u}{\partial C_{nm}}\right| + \left|\frac{\partial u}{\partial C_{n0}}\right|\right) \\ &= \frac{4\left|dC_{n0}\right|}{H_{nm}}\left(\left|C_{nm} - \frac{1}{2}C_{n0}\right| + \left|S_{nm} - \frac{1}{2}C_{n0}\right| + \frac{1}{2}\left|S_{nm} - C_{nm}\right|\right) \end{aligned} \tag{30}$$

En utilisant la relation (6) on obtiend finalement

$$|dA_{nm}| \leq \frac{6}{\sqrt{H_{nm}}} \left| \frac{dC_{n0}}{C_{n0}} \right| \tag{31}$$

Ainsi l'expression pour la borne supérieure de l'erreur d'évaluation des CHCA est donnée par la relation suivante :

$$\left|dA^{m_0}{}_{nm}\right| \leq 6\left(\frac{1}{\sqrt{H_{nm}}} + \frac{m}{m_0\sqrt{H_{nm_0}}} \right) \left| \frac{dC_{n0}}{C_{n0}} \right| \tag{32}$$

À prime abord cette dernière relation semble aussi catastrophique que celle se rapportant à l'erreur sur les CHC. Toutefois à la section sur les résultats optiques nous allons voir que l'évaluation des CHCA semble significativement moins critique que celle des CHC.

2.2.2 Incertitudes dues à un mauvais ajustement de l'orientation des rosettes complémentaires

La mise en oeuvre optique de l'extraction des CHC est basée sur la relation (7). Ceci implique que pour chacun des ordres harmoniques m, une paire de rosettes doit être utilisée pour l'évaluation du C_{nm} et du S_{nm} impliqués dans la relation (7). Ces rosettes sont complémentaires l'une à l'autre et l'orientation de l'objet relativement aux rosettes doit être identique pour les deux rosettes complémentaires, sans quoi la relation (7) ne peut plus s'appliquer. Il est possible de se faire une idée de la précision avec laquelle doit se faire l'ajustement de l'orientation des rosettes. Pour cela les relations (18) et (19) sont d'une grande utilité.

Supposons qu'il y ait une imprécision dans l'orientation des rosettes d'ordre m_0. Soit α, l'angle que font entre eux les axes des rosettes complémentaires d'ordre m_0. On peut considérer que l'orientation de l'objet relativement aux rosettes en *cos* et en *sin* est spécifée respectivement par les angles ϕ_0-α/2 et ϕ_0+α/2, ϕ_0 étant l'angle entre une certaine direction préférentielle sur l'objet et l'un des axes de coordonnées dans le référentiel des rosettes, c'est-à-dire $\phi_0 = \phi + (1/m_0)\,\phi_{nm_0}$. Selon la relation (7), l'estimation de la valeur du CHC est donnée par la relation suivante :

$$\widehat{H}_{nm}(f) = \frac{4}{C_{n0}^2}\left(\left[C_{nm}\left(f;\phi_0-\frac{\alpha}{2}\right)-\frac{1}{2}C_{n0}\right]^2+\left[S_{nm}\left(f;\phi_0+\frac{\alpha}{2}\right)-\frac{1}{2}C_{n0}\right]^2\right) . \quad (33)$$

Selon les relations (18) et (19), cette expression se ramène à :

$$\widehat{H}_{nm}(f) = \frac{4|F_{nm}|^2}{C_{n0}^2}\left(\left[\cos\left(m\left(\phi_0-\frac{\alpha}{2}\right)\right)\right]^2+\left[\sin\left(m\left(\phi_0+\frac{\alpha}{2}\right)\right)\right]^2\right). \quad (34)$$

En développant les cosinus et sinus dans la dernière expression, on obtient :

$$\widehat{H}_{nm}(f) = \frac{4|F_{nm}|^2}{C_{n0}^2}\left[1+\sin(2m\phi_0)\sin(m\alpha)\right] . \quad (35)$$

Puisque $H_{nm}(f) = 4|F_{nm}|^2/C_{n0}^2$, on obtient finalement l'expression pour l'incertitude relative

$$\frac{E}{H_{nm}(f)} \equiv \left|\frac{\widehat{H}_{nm}(f)-H_{nm}(f)}{H_{nm}(f)}\right| = \left|\sin(2m\phi_0)\sin(m\alpha)\right| \le \left|\sin(m\alpha)\right| \sim m\left|\alpha_{radian}\right| . \quad (36)$$

Dans cette dernière relation l'approximation est valide seulement lorsque $m\alpha_{radian}$ est petit. Selon l'orientation de l'objet, c'est-à-dire selon la valeur de ϕ_0, l'incertitude relative due à une mauvaise mise en correspondance angulaire des rosettes peut être aussi grande que $|sin(m\alpha)|$. À titre d'exemple considérons une erreur d'orientation de 1° pour la paire de rosettes de l'ordre harmonique 10. Dans ce cas $|sin(m\alpha)| = sin(10°) = 0.17$, soit une incertitude relative de 17%. De façon à pouvoir utiliser la relation (7), Il est important que les rosettes complémentaires soient mises en correspondance angulaire l'une par rapport à l'autre avec une certaine précision. Toutefois il y a mieux à faire puisque l'on peut dériver une formulation plus générale que la relation (7), laquelle permet de tenir compte des erreurs de mise en correspondance des rosettes. Ceci est fait dans la prochaine sous-section . Les calculs homologues dans le cas des CHCA n'ont pu être effectuées en raison de la complexité des équations

2.3 Généralisation de la formulation des CHC optiques

Il a été établi à la section 2.2, que les erreurs dues à la mauvaise mise en correspondance angulaire des rosettes complémentaires, ainsi que celles dues aux différences de gain et de biais dans les deux canaux de mesures, génèrent des incertitudes importantes sur les CHC. Dans cette section la question de la reformulation des CHC optiques (relation 7) est abordée, afin de tenir compte de telles erreurs.

Une expression plus générale pour le calcul des CHC est donc dérivée à la sous-section 2.3.1. Elle est fonction de cinq paramètres qui dépendent des conditions expérimentales. Une méthode numérique de calibration des paramètres de ce modèle est ensuite présentée à la section 2.3.2.

2.3.1 Reformulation du calcul des CHC

Les expressions (18) et (19) expriment la dépendance théorique des mesurables en fonction de l'angle de rotation de l'objet par rapport aux rosettes. En pratique toutefois, les conditions réelles de mesure introduisent des écarts que nous modéliserons par les relations suivantes:

$$C_{nm}(f;\phi-\alpha/2,A_c,B_c)-\frac{1}{2}C_{n0}(f,\phi) = A_c|F_{nm}|\cos(m[\phi_0-\alpha/2])+B_c \qquad (37)$$

et

$$S_{nm}(f;\phi+\alpha/2,A_s,B_s)-\frac{1}{2}C_{n0}(f,\phi) = A_s|F_{nm}|\sin(m[\phi_0+\alpha/2])+B_s \qquad (38)$$

Ce modèle rend compte des différences de gains (A_c/A_s) et de biais (B_c-B_s) dans les deux canaux de mesures, ainsi que d'une erreur relative de mise en correspondance angulaire des rosettes d'angle α. Seule subsiste la supposition que les centres des rosettes sont parfaitement alignés sur le même point de l'objet.

Lorsque les relations (18) et (19) sont exprimées en fonction des relations (37) et (38), une formulation plus générale de la relation (7) est obtenue:

$$\begin{aligned} H_{nm}(\alpha,A_c,A_s,B_c,B_s) = \left(\frac{2}{C_{n0}A_cA_s\cos(m\alpha)}\right)^2 \Big[A_s^2\Big(U_{nm}^c - B_c\Big)^2 \\ -2A_cA_s\Big(U_{nm}^c - B_c\Big)\Big(U_{nm}^s - B_s\Big)\sin(m\alpha) + A_c^2\Big(U_{nm}^s - B_s\Big)\Big] \end{aligned}, \quad (39)$$

où

$$\begin{aligned} U_{nm}^c(f;\phi_0-\alpha/2,A_c,B_c) &= C_{nm}(f;\phi_0-\alpha/2,A_c,B_c) - \frac{1}{2}C_{n0}(f,\phi_0) \\ U_{nm}^s(f;\phi_0+\alpha/2,A_s,B_s) &= S_{nm}(f;\phi_0+\alpha/2,A_s,B_s) - \frac{1}{2}C_{n0}(f,\phi_0) \end{aligned} \quad (40)$$

représentent les termes associés aux mesures. La relation (39) intègre donc les compensations nécessaires au calcul des CHC lorsque les paramètres α,A_c,A_s,B_c,B_s sont connus. À la section suivante, on présente une technique d'optimisation qui permet d'estimer ces paramètres afin de caractériser le système optique.

2.3.2 Caractérisation du système optique

Afin de déterminer les paramètres de compensation optimaux pour chacun des canaux, un objet de référence (f_r) est utilisé. Il peut s'agir de n'importe quel objet, pourvu que sa signature CHC $H_{nm}(f_r)$ soit connue pour tous les ordres considérés (par simulation numérique par exemple). L'objet de référence est successivement présenté au système optique, lors d'une série de rotations d'angles ϕ_i, $i \in [1,N]$. Les valeurs des canaux $C_{nm}(f_r;\phi_i)$, $S_{nm}(f_r;\phi_i)$ et $C_{n0}(f_r;\phi_i)$ sont alors recueillies. A partir de ces mesures, les valeurs $U_{nm}^c(f_r;\phi_i)$ et $U_{nm}^s(f_r;\phi_i)$ de la relation (40) sont calculées. Selon les relations (37) et (38), elles décrivent idéalement un cosinus et un sinus en fonction de l'angle ϕ_i.

Par la méthode d'optimisation aux moindres carrés de Levenberg-Marquardt [10], un cosinus et un sinus sont ajustés respectivement sur les courbes décrites par U_{nm}^c et U_{nm}^s, tels que:

$$\begin{aligned} \widehat{U}_{nm}^c(\phi_0) &= a_c\cos(m[\phi_0-\delta_c]) + b_c \\ \widehat{U}_{nm}^s(\phi_0) &= a_s\sin(m[\phi_0-\delta_s]) + b_s \end{aligned}, \quad (41)$$

où l'on s'assure que $a_x = |a_x|$ en absorbant, le cas échéant, le signe négatif dans la phase ($\delta_x \to \delta_x + \pi/m$). Ensuite, on prend la peine de rapporter les phases dans l'intervalle $\delta_x \in [0,2\pi]$.

La comparaison des relations (37) et (38) avec la relation (41) permet de déduire que $B_c = b_c$ et $B_s = b_s$ et que le paramètre α, l'erreur de mise en correspondance angulaire des rosettes, est obtenu par $\alpha = \delta_s - \delta_c$.

Pour la détermination des paramètres A_c et A_s, on trouve:

$$A_x = \frac{a_x}{|F_{nm}|}. \tag{42}$$

Or la valeur de $|F_{nm}|$ n'est pas connue *a priori*, mais plutôt estimée à partir de la valeur du CHC de référence et de la relation $H_{nm}(f_r) = 4|F_{nm}|^2 / C_{n0}^2$. Puisque l'évaluation de $C_{n0}(\phi_i)$ est expérimentale, on trouve l'estimée la plus probable pour $|F_{nm}|$ au sens des moindres carrés, en prenant la moyenne:

$$\widehat{|F_{nm}|} = \frac{[H_{nm}(f_r)]^{1/2} \langle C_{n0}(\phi_i) \rangle}{2}. \tag{43}$$

En remplaçant dans la relation (42), on trouve enfin:

$$A_x = \frac{2a_x}{[H_{nm}(f_r)]^{1/2} \langle C_{n0}(\phi_i) \rangle}. \tag{44}$$

Cette méthode permet donc l'évaluation expérimentale des cinq paramètres considérés dans le modèle de la relation (39), réalisant ainsi la caractérisation du système optique. Les paramètres ainsi déterminés sont estimé une seule fois et sont utilisés par la suite pour le calcul des CHC à l'aide de la relation (39).

CHAPITRE 3

Systèmes optiques pour l'extraction des CHC

Pour réaliser l'extraction des CHC, deux approches optiques ont été mises à l'essai. La première approche met à profit la capacité des réseaux de diffraction à produire des répliques de la figure d'illumination. Il s'agit d'une architecture à canaux multiples avec une illumination cohérente. L'objet segmenté est affiché sur un afficheur à cristaux liquides (ACL) tandis que les rosettes sont affichées sur des masques en chrome. La seconde méthode implique deux ACL, un premier pour l'affichage de l'objet et l'autre pour celui des rosettes. Une source incohérente est utilisée pour l'illumination du ACL de l'objet et celui-ci est imagé sur le second ACL. L'approche avec réseau de diffraction permet l'évaluation de plusieurs CHC simultanément toutefois sa réalisation implique un éclairage cohérent (laser) et les questions de bruit qui y sont associés. La méthode à deux ACL ne permet pas l'évaluation simultanée d'un nombre important de CHC dans la version présentée mais elle offre l'avantage de permettre une meilleure qualité au niveau de l'optique.

3.1 Système à réseau de diffraction

Le système à réseau de diffraction est celui des deux systèmes étudiés qui est le plus complexe à réaliser. Il comprend de nombreux composants qu'il faut ajuster avec précision. Toutefois, il présente certains avantages sur l'approche à deux ACL puisque toutes les mesures sont prises à l'intérieur de la période de temps correspondant à un seul cycle vidéo (1/30 seconde). Ceci permet des simplifications au niveau du post-traitement et de la synchronisation des processus.

La description du principe de fonctionnement du système à réseau de diffraction et des principales étapes menant à sa réalisation est présentée dans les prochaines sections.

3.1.1 Principe de fonctionnement

Il est bien connu que les réseaux de diffraction utilisés en spectroscopie produisent généralement non pas un spectre mais en fait plusieurs répliques du spectre [11]. C'est une conséquence généralement non souhaitée de la structure périodique créée dans le but de décomposer le spectre selon les longueurs d'onde. La forme de la cellule élémentaire, qui est répétée un grand nombre de fois pour créer le réseau, a une grande influence sur la répartition de l'énergie selon les différentes répliques ou ordres[1] du spectre. La cellule élémentaire des réseaux spectroscopiques est optimisée de façon à concentrer l'énergie sur une seule des répliques du spectre.

Dans un spectromètre le réseau de diffraction est associé à un système d'imagerie [11]. Le système d'imagerie a pour fonction de faire l'image de la source, en l'occurence une fente plus ou moins étroite, dans le plan de sortie de l'appareil. Lorsque la fente d'entrée est éclairée par une lumière monochromatique, provenant d'une source laser ou d'une lampe incandescente munie d'un filtre, on obtient dans le plan de sortie plusieurs répliques de la fente d'entrée. Le système d'extraction des CHC fonctionne sur ce même principe. À la place de la fente on place l'objet affiché sur un ACL et celui-ci est éclairé par un faisceau laser. La figure 3.1 illustre ce principe. Pour évaluer les CHC, les rosettes sont placées dans le plan de sortie et la lumière transmise est collectée sur des détecteurs.

1. On utilise généralement *ordres de diffraction* ou tout simplement *ordres* comme terme descriptif des différentes répliques du spectre.

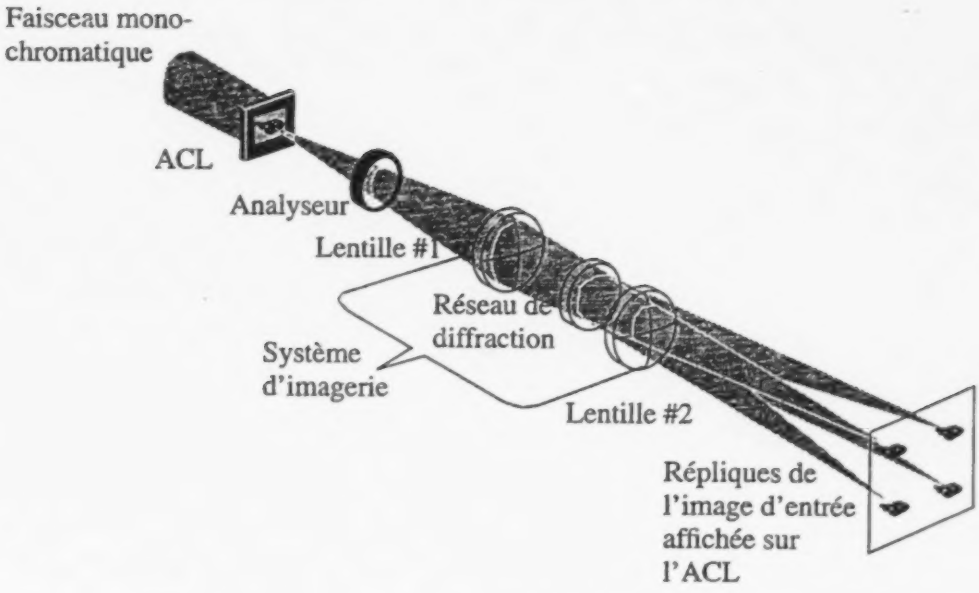


FIGURE 3.1 : Principe de fonctionnement du système à réseau de diffraction.

Dans ce système on utilise un réseau de diffraction d'un type spécial. Il s'agit d'un réseau de phase[1] dont la structure périodique est bidirectionnelle, c'est-à-dire que la microtopographie du réseau est périodique selon deux directions orthogonales. La structure bidirectionnelle du réseau fait en sorte qu'il agit comme deux réseaux simples en cascade et orientés à 90° l'un de l'autre. Ceci permet donc de générer une mosaïque de répliques de l'objet affiché à l'entrée du système. De plus la cellule élémentaire du réseau est optimisée de manière à ce que l'énergie soit concentrée et répartie uniformément sur les 7 x 7 premiers ordres.

Pour opérer le système il faut une source monochromatique ou à spectre très étroit. L'utilisation d'une source à spectre large n'est pas acceptable puisque la décomposition

1. Un réseau de phase doit sont action sur le faisceau incident uniquement par l'introduction de variations localisées du parcourt optique. La longueur des variations du parcourt optique est comparable à la longueur d'onde du faisceau incident.

spectrale produite par le réseau amènerait une forte dégradation des images. La contrainte de la monochromaticité entraîne l'obligation d'utiliser une source laser. Il serait difficile de faire autrement puisque les autres types de source à spectre suffisamment étroit ne sont pas très intenses et nécessiteraient l'utilisation de détecteurs très sensibles (photomultiplicateurs par exemple).

Sur le schéma de la figure 3.1 quelques composants indispensables ont été omis. Ceci a été fait délibérément dans le but de mieux faire ressortir le parallèle entre le système et un spectromètre. La figure 3.2 complète le schéma de la figure 3.1.

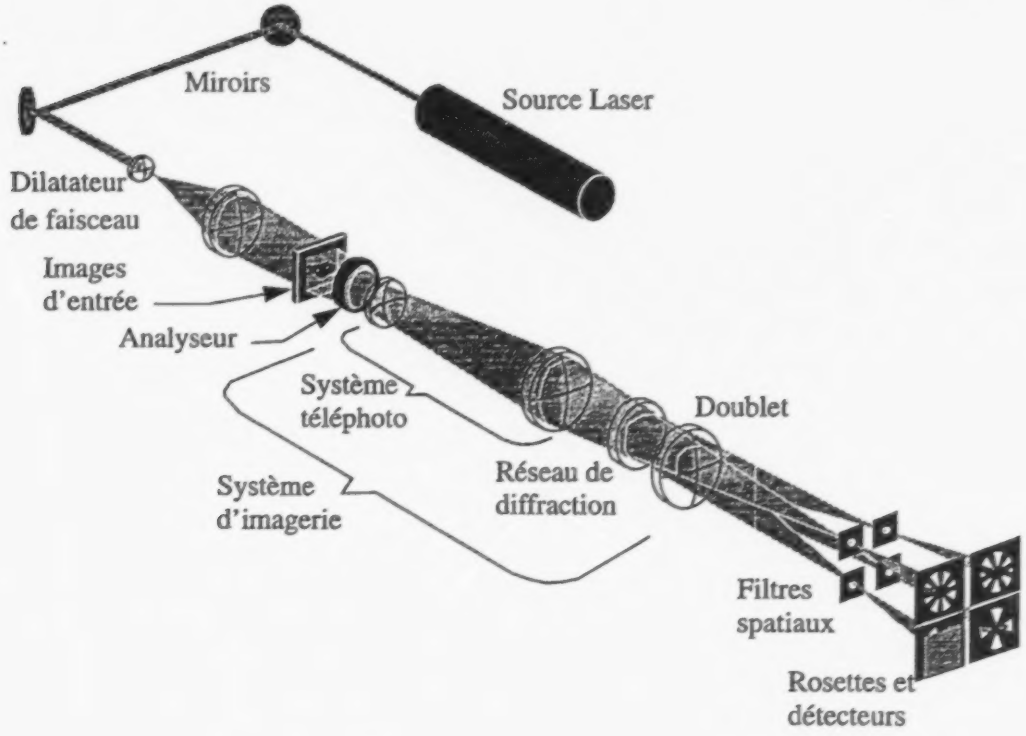


FIGURE 3.2 : Schéma complet du système à réseau de diffraction.

3.1.2 Construction du montage à réseau de diffraction

Le montage à réseau de diffraction comprend plusieurs petites composantes optiques qui doivent être positionnés avec précision. L'architecture parallèle du système laisse peu de place pour certaines de ces composantes. Malgré l'espace exigu, des moyens simples ont dû être mis de l'avant pour assembler le système. Certaines composantes ont due être fabriquées sur mesure, spécifiquement pour la réalisation du système. C'est le cas des rosettes et de certaines pièces d'ajustement mécanique.

Dans la prochaine section, nous allons donner une brève description des principales composantes entrant dans la réalisation du système. Celle-ci sera suivie d'une section traitant des différents aspects liés à l'assemblage du système.

3.1.2.1 Principaux composants du montage

Comme le montre la figure 3.2, les principaux composants sont le système d'illumination monochromatique, l'afficheur à cristaux liquides, le système optique d'imagerie, le réseau de diffraction, la matrice des filtres spatiaux, la matrice des rosettes et la matrice des photodétecteurs.

Système d'illumination monochromatique :

Il est consitué d'une source laser He-Ne de 15 mW. Le faisceau est redirigé par réflexion sur deux miroirs à orientation ajustable. Le faisceau laser est dilaté à l'aide d'un téléscope de type Képler tout en étant filtré spatialement par un sténopé placé au foyer interne du téléscope. La distance entre les deux lentilles du téléscope est ajustée de façon à produire un faisceau qui converge lentement.

Afficheur à cristaux liquides (ACL) :

Pour cette application il a fallu choisir un afficheur de bonne qualité ayant des pixels de petites dimensions. Un ACL de marque Hitachi a été retenu. Il s'agit de l'un des meilleurs ACL disponibles actuellement sur le marché. Les temps de réponse de cet ACL sont respectivement de 45 ms pour la monté et de 15 ms pour la descente. Le contraste pour un éclairement de faible ouverture numérique à incidence normale est supérieur à 1000:1. Le circuit électronique permettant l'opération de l'ACL à partir d'un signal VGA est très compact. L'électronique d'adressage et le ACL proviennent en fait d'un appareil de réalité virtuelle commercialisé par la firme *Seattle Sight System*. Sa compacité et le fait que l'on puisse réellement contrôler chaque pixel individuellement (pas d'échantillonnage) sont des atouts importants pour la présente application. Le tableau 1 regroupe les principales caractéristiques de l'ACL.

TABLEAU 1. Caractéristiques de l'ACL Hitachi/Seatlle Sight System

Distance interpixel	32 µm x 32 µm
Résolution	480 x 640
Dimension de la surface active	15.4 mm x 20.4 mm
Contraste[a]	> 1000:1
Temps de monté	45 ms
Temps de descente	15 ms
Dimension hors tout du ACL	27 mm x 21mm x 1 mm
Distorsion des surfaces[b]	~ 2 λ @ 633 nm
Dimension hors tout de l'électronique d'adressage	165 mm x 45 mm x 10 mm

a. Mesuré à incidence normale avec un laser He-Ne (633 nm) collimé.

b. Mesuré à l'aide d'un interféromètre de Fizeau.

Système d'imagerie :

Le système d'imagerie sert à produire l'image de l'ACL dans le plan des rosettes. Le système d'imagerie est divisé en deux parties. La première est un système téléphoto inverse permettant de réduire la longueur focale effective dans un format relativement compact. La

seconde partie comprend un seul doublet de 750 mm de focale. La longueur focale de cette dernière lentille est imposée par le pas du réseau de diffraction et par l'espacement désiré entre les répliques des images dans le plan des rosettes. La longueur focale effective du système téléphoto est d'environ 780 mm, ce qui donne un facteur de grandissement d'environ 1.05 pour le système global. Le schéma du système d'imagerie apparait à la figure 3.3 et les données le concernant sont réunies dans le tableau 2.

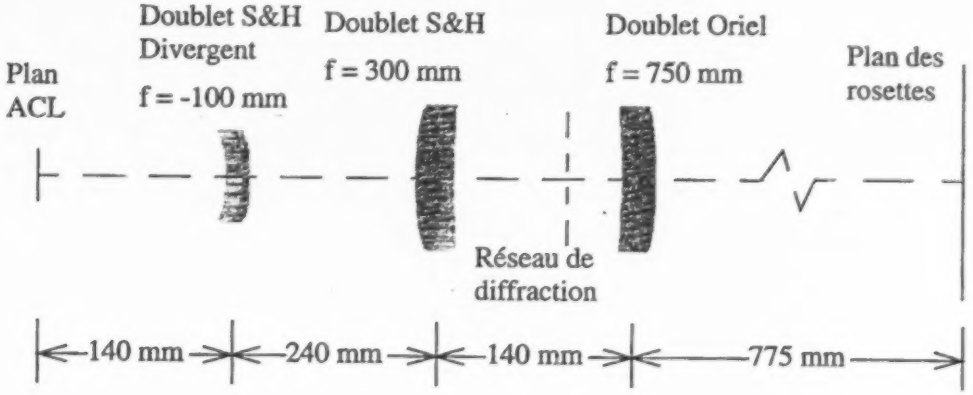


FIGURE 3.3 : Système d'imagerie du système à réseau de diffraction. (S&H = Spinler and Hoyer)

TABLEAU 2. Données se rapportant au système d'imagerie

Composant	Focale (mm)	Distance[a] (mm)
ACL	N/A	0
Doublet divergent Spinler & Hoyer # 325222	- 100	140
Doublet Spinler & Hoyer # 322305	300	240
Doublet Oriel # 42685	750	140
Plan des rosettes	N/A	775

a. Mesurée relativement au composant précédent

Réseau de diffraction :

Le réseau de diffraction utilisé dans le montage est de type Dammann binaire. Il a été fabriqué par l'INO spécifiquement pour ce type d'application. Il s'agit d'un réseau de phase comportant un microrelief 'gravé' dans de la photorésine. La microtopographie du réseau comporte une périodicité selon deux axes orthogonaux et les reliefs sont en palier avec seulement deux niveaux (réseau de phase binaire). La figure 3.4 donne un aperçu de la microtopographie d'un réseau Dammann binaire bidirectionnel.

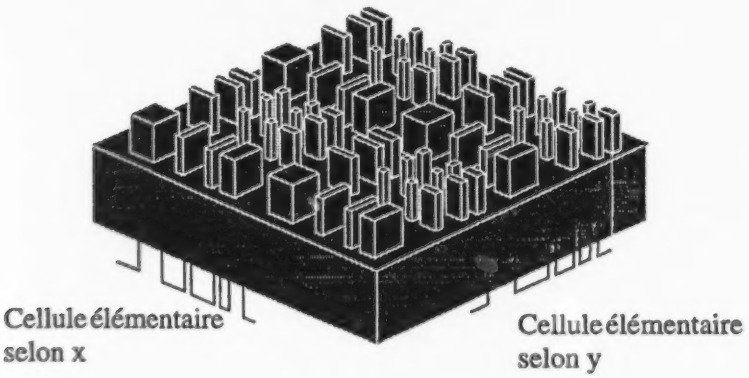


FIGURE 3.4 : Microtopographie type d'un réseau de Dammann binaire bidirectionnel.

La cellule élémentaire de ce réseau est optimisée pour concentrer l'énergie dans les 7 x 7 premiers ordres de diffraction. La disparité de la répartition de l'énergie entre les 7 x 7 ordres centraux est inférieure à 20%. Le format de la surface disponible est de 50 mm x 50 mm.

Rosettes :

Ce sont les composants qui ont nécessité le plus d'efforts à fabriquer. On a tout d'abord tenté de les produire en utilisant un appareil qui permet d'écrire par balayage laser sur une pellicule photographique (BLPP). Plusieurs essais ont été necéssaires pour atteindre les

performances optimales accessibles par la technologie disponible. Ceci n'a pu être obtenu qu'en créant directement les fichiers 'postcript'[1] sans intermédiaire de façon à pouvoir contrôler tous les paramètres. Il a fallu également élaborer une procédure de calibrage des tons de gris, dans le but de compenser pour la non-linéarité de l'écriture. Le calibrage des tons de gris nécéssite la génération d'un figure de référence contenant 256 zones de tons de gris 'postcript' distincts. La transmittance de chacun de ces 256 tons de gris 'postcript' imprimé est individuellement mesurée à l'aide d'un système microscopique spécialement assemblé pour cette tâche. Les mesures ainsi obtenues permettent de caractériser la réponse de l'appareil d'impression BLPP. Connaissant la réponse du système, il est facile de créer une table de correspondance permettant la compensation de la non-linéarité. Ceci se fait au dépend de la plage dynamique puisque la compensation implique l'élimination de nombreux tons de gris 'postcript'. Sur les 256 tons de gris postcript disponibles seuls une centaine sont utilisés par la table de correspondance. Malgré tous les efforts déployés, la technologie BLPP n'a pu permettre d'atteindre une résolution suffisante et une qualité optique satisfaisante.

Dans ces circonstances, la seconde approche considérée a été l'*écriture par faisceau d'électrons* (EFE) sur masque en chrome. Cette technologie permet la création de masques binaires d'amplitude (deux tons de gris, transmission ~ 0 / transmission maximale) dont la taille des pixels est de 1 µm x 1 µm. En groupant les pixels en superpixels de 10 µm x 10 µm (100 pixels) on peut créer des figures comportant jusqu'à 100 tons de gris. Ainsi la technologie EFE assure des résolutions supérieures par un facteur 4 à celles accessibles par procédé BLPP. De plus la qualité optique des masques en chrome obtenus par EFE est excellente. Toutefois la technologie EFE est beaucoup plus coûteuse et a impliqué des délais beaucoup plus important en comparaison au procédé BLPP. Malheureusement, les délais de livraison des masques en chrome ont été de plusieurs mois, et ont retardé considérablement le déroulement du projet.

La figure 3.5 montre la rosette d'ordre 2 obtenue respectivement par la technologie BLPP à droite et par la technologie EFE à gauche. On constate qualitativement que la

1. *Postscript* est une marque de commerce de Adobe Systems Incorporated. L'appellation désigne un langage informatique interprété, doté de puissantes fonctions graphiques. Le language *Postscript*, indépendants de l'appareil de rendu, constitue un standard.

technologie EFE est nettement supérieure et qu'elle permet une bonne résolution dans le centre de la rosette.

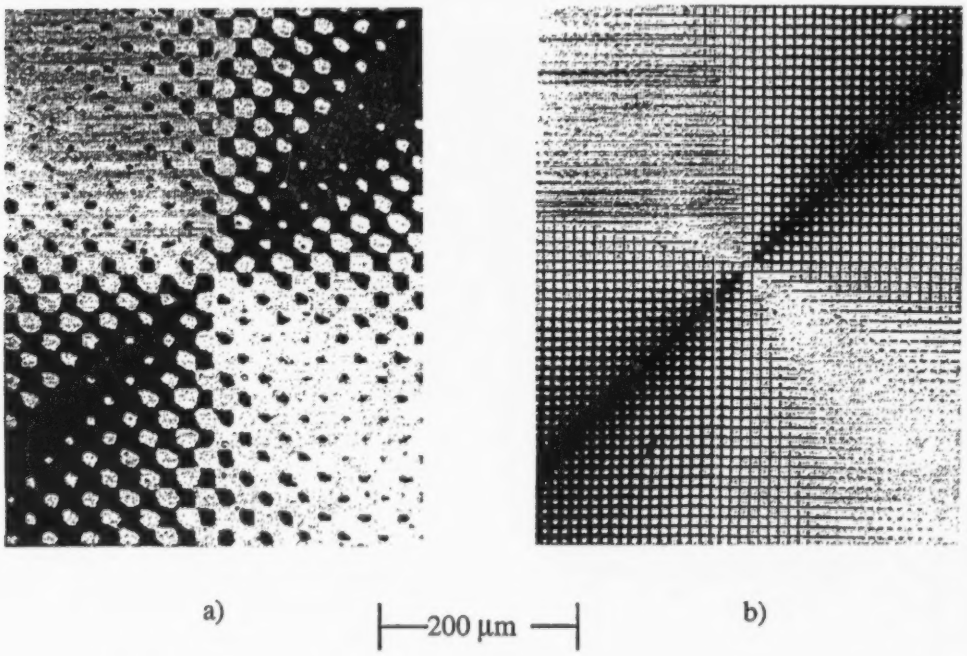


FIGURE 3.5 : Microphotographie de la partie centrale de la rosette d'ordre 2 produite en a) par procédé BLPP et en b) par écriture par faisceau d'électrons. Le facteur de grandissement est d'environ 145 X.

Filtres spatiaux :

Les filtres spatiaux sont de simples plaques opaques transpercées d'un trou de 4 mm de diamètre. Il y a un tel filtre pour chacune des rosettes. Les filtres sont placés dans le plan qui est conjugué à celui du sténopé du système d'illumination (source). On retrouve dans ce plan les répliques de la TF de l'image affichée par l'ACL. Aux répliques générées par le réseau de diffraction sont superposées celles générées par la "pixellisation" de l'ACL. Pour chacune des répliques générées par le réseau, un filtre spatial est utilisé pour éliminer toutes les répliques produites par l'ACL à l'exception de l'ordre centrale. Ceci est illustré à la figure

3.6. L'utilisation des filtres spatiaux permet d'éliminer virtuellement toute trace de l'effet de la pixellisation de l'ACL.

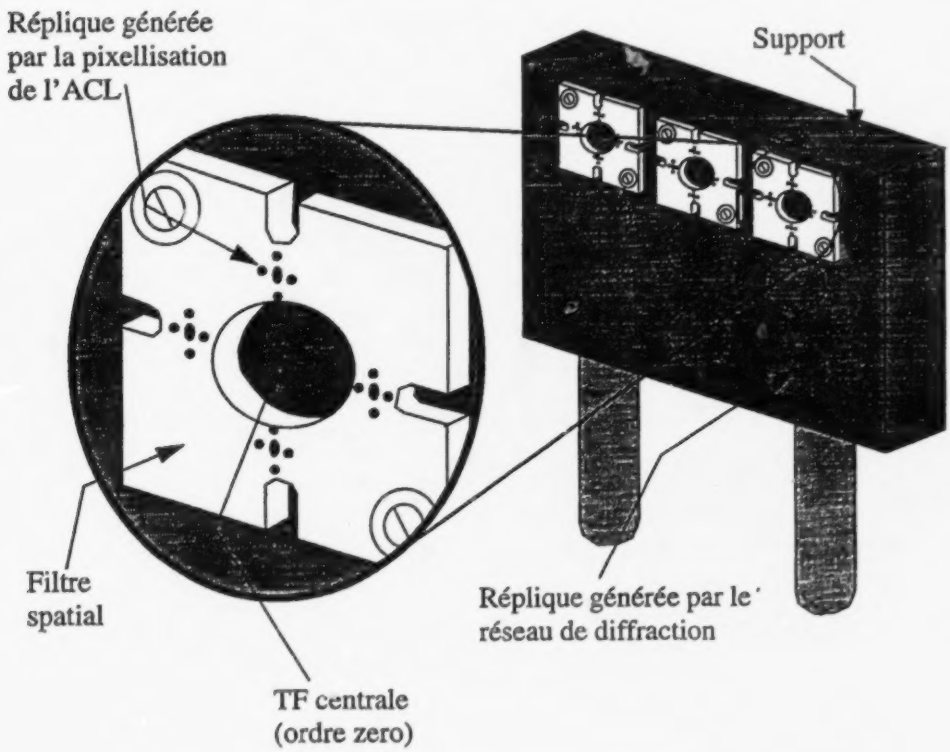


FIGURE 3.6 : Filtres spatiaux sur leur support et les TF.

Détecteurs :

Des photodiodes de dimensions légèrement plus grandes que celles des rosettes ont été utilisés. Les détecteurs sont placés immédiatement en arrière des rosettes. Ceci permet l'interception par les détecteurs de tous les rayons de lumière transmis par les rosettes, sans devoir recourir à de l'optique supplémentaire entre les rosettes et les détecteurs. Le tableau 3 donne les principales caractéristiques des détecteurs utilisés.

TABLEAU 3. Principales caractéristiques des détecteurs.[a]

Fabriquant	EG&G Vartec
Type	VTS 30 82
Dimension de la surface active	~ 9 mm x 10 mm
Sensibilité maximale	0.55 A/W
Dimensions globales	~ 10 mm x 10 mm x 2 mm

a. Données venant du manuel du fabriquant.

L'amplification des signaux provenant des détecteurs se fait à l'aide du circuit illustré à la figure 3.7. Il comprend deux étages d'amplification. Le premier comprend un amplificateur opérationel CA3140AE qui offre à la fois beaucoup de gain et de stabilité. Le premier étage est placé très près du détecteur de façon à réduire les bruits parasites collectés par les fils du détecteur. Ce premier amplificateur opérationnel de concert avec la résistance R_1 sert à convertir le courant de la photodiode en une tension.

La valeur de la tension de sortie du premier amplificateur opérationnel est environ 10 millions de fois celle du courant de la photodiode. La valeur de la résistance R_1 est très élevée de façon à donner le maximum de gain possible au départ. Ceci est une condition nécessaire pour obtenir le meilleur rapport signal à bruit possible. Pour réduire d'avantage le bruit le condensateur C_1 a été ajouté. La capacitance est choisie de façon à ce que le circuit d'amplification ne laisse pas passer les signaux à haute fréquence, c'est-à-dire le bruit.

La résistance variable R_3 sert à l'ajustement du zéro, lorsque le détecteur n'est pas exposé à la lumière. La résistance R_2 sert pour compenser les courants de polarisation de l'amplificateur opérationnel. C_2, C_3, C_4 et C_5 servent de condensateurs de découplage pour filtrer les alimentations.

Le second étage d'amplification comprend un amplificateur opérationnel OP27G qui offre une grande précision et génère très peu de bruit. Ce second étage d'amplification à un gain variable de 1 à 101. L'ajustement du gain se fait grâce à la résistance variable R_4. La

diode Zener D_2 est placée à la sortie du circuit et a pour fonction de protéger les entrées de la carte d'acquisition de données. Cette diode limite à 10 V la tension maximale de sortie du circuit d'amplification. Les entrées de la carte d'acquisition peuvent supporter des tensions ne dépassant pas 11V.

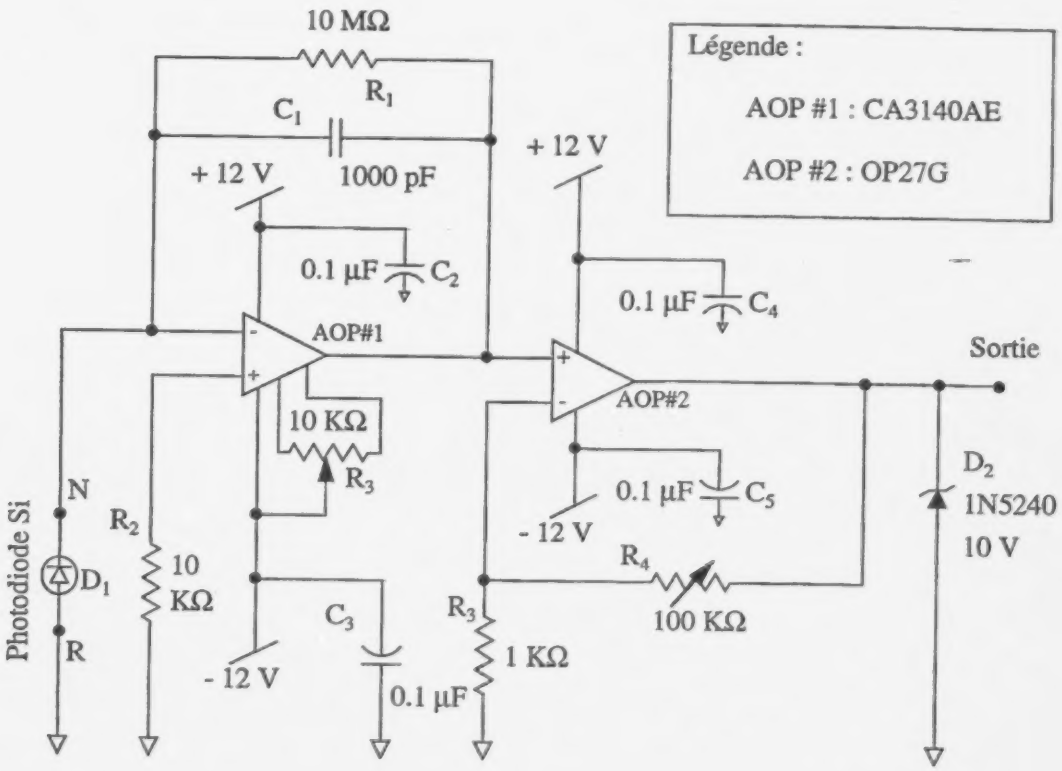


FIGURE 3.7 : Schéma du circuit d'amplification des photodétecteurs.

3.1.3 Assemblage du montage

L'assemblage du montage est une opération qui a demandé beaucoup de temps en raison de la précision requise pour l'ajustement de certaines composantes. L'architecture

parallèle du montage complique grandement la situation. Pour permettre l'ajustement précis des rosettes et des filtres spatiaux, des dispositifs mécaniques spéciaux ont du être construits. Les prochaines sections portent sur les différents aspects liés à l'assemblage du montage. L'information y est résumée et le lecteur désireux de prendre connaissance des détails peut le faire en consultant l'annexe A.

3.1.3.1 Mécanique

Pour permettre le positionnement individuel des rosettes avec une grande précision en tenant compte de l'exéguité du montage, il a fallu concevoir et construire un dispositif mécanique de précision. Il s'agit d'un mécanisme amovible permettant l'ajustement d'une seule rosette à la fois. La figure 3.8 montre ce mécanisme de même que le support des rosettes. Le mécanisme de positionnement des rosettes se fixe en différents endroits sur le support des rosettes. Le mécanisme de positionnement permet le déplacement latérale et verticale sur une plage de ± 1 mm et des rotations de ± 5°.

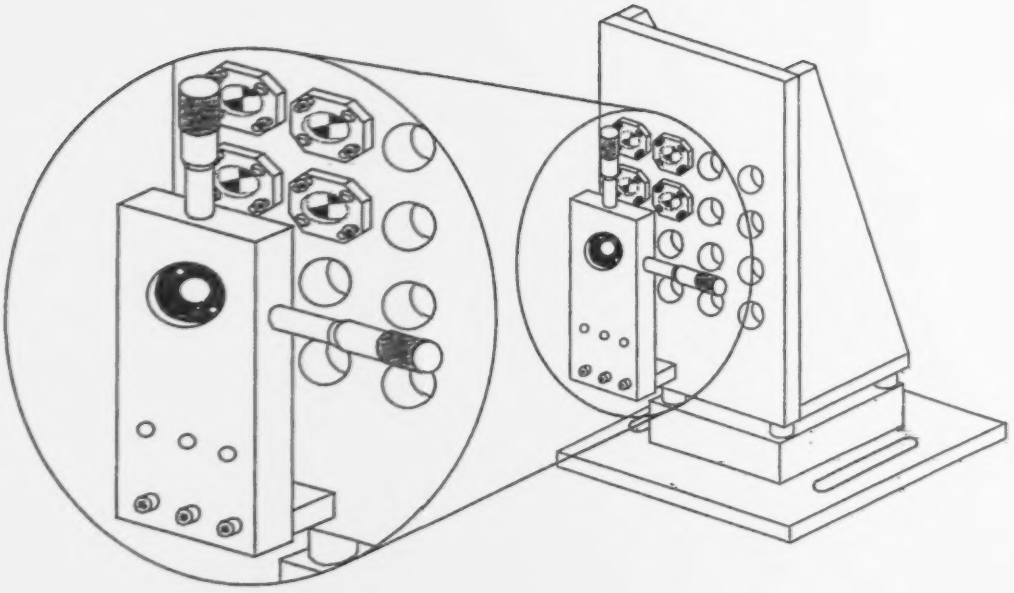


FIGURE 3.8 : Mécanisme de positionnement et support de fixation des rosettes.

Les rosettes sont montées dans des supports individuels. Ces supports sont dotés de deux concavités par où vient s'agripper le mécanisme de positionnement. Les rosettes doivent être bien centrées et bien orientées dans leur support respectif. Ceci est necéssaire puisque la plage de déplacement et de rotation accessible par le mécanisme de positionnement sont relativement limitées. Les ajustements des rosettes dans leur support individuel a nécessité un montage spécifique. Les rosettes ont été collées dans leur support avec une erreur de centrage inférieure à 200 µm et une erreur d'orientation ne dépassant pas 0.5°.

Le dispositif de centrage des filtres spatiaux est bien plus simple que celui des rosettes. Le dispositif de centrage des filtres est illustré à la figure 3.9. Chacun des filtres comprend quatre fentes. Ces fentes servent d'appuis pour les outils de déplacement. Tel qu'illustré à la figure 3.9 le support des filtres est une simple plaque doté de grande ouverture par où passe les rayons de lumière après le filtrage spatiale. La plaque comprend également de nombreux

petits trous par où se glisse le pivot des outils d'ajustement. L'ajustement des rosettes et des filtres spatiaux se fait à l'aide d'un microscope équipé d'une caméra vidéo.

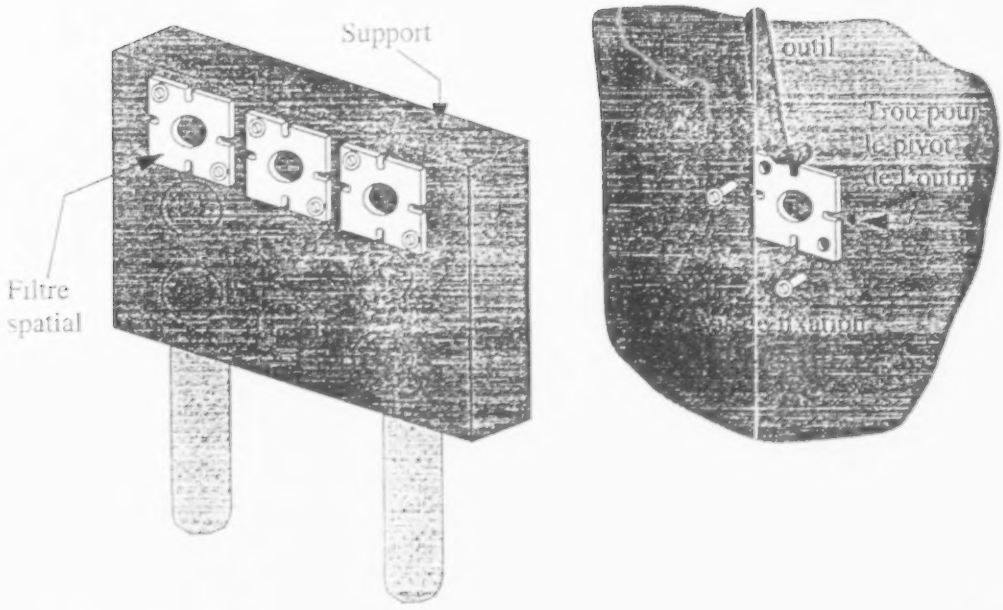


FIGURE 3.9 : Filtres spatiaux, leurs supports et l'outil de déplacement.

3.1.3.2 Alignement

L'alignement est l'opération qui consiste à positionner et à orienter correctement les composants optiques les uns par rapport aux autres. L'alignement vise à faire coincider les axes optiques de tous les composants et à disposer les composants selon les espacements requis. L'alignement du montage à réseau de diffraction est une opération relativement complexe en raison du nombre important de composants impliqués (degrés de liberté) et de la précision relativement élevée requise. L'alignement se fait en plusieurs étapes et débute par un préalignement destiné à établir les positions approximatives des composants sur la table

petits trous par où se glisse le pivot des outils d'ajustement. L'ajustement des rosettes et des filtres spatiaux se fait à l'aide d'un microscope équipé d'une caméra vidéo.

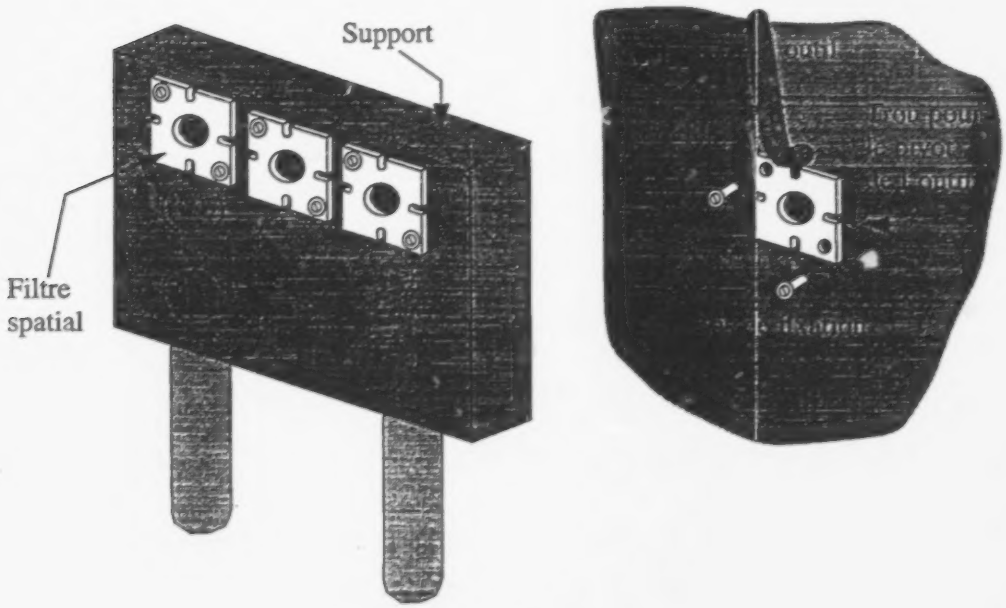


FIGURE 3.9 : Filtres spatiaux, leurs supports et l'outil de déplacement.

3.1.3.2 Alignement

L'alignement est l'opération qui consiste à positionner et à orienter correctement les composants optiques les uns par rapport aux autres. L'alignement vise à faire coincider les axes optiques de tous les composants et à disposer les composants selon les espacements requis. L'alignement du montage à réseau de diffraction est une opération relativement complexe en raison du nombre important de composants impliqués (degrés de liberté) et de la précision relativement élevée requise. L'alignement se fait en plusieurs étapes et débute par un préalignement destiné à établir les positions approximatives des composants sur la table

optique. Les prochains paragraphes décrivent sommairement les étapes plus critiques. De plus amples détails sont fournis à l'annexe A.

L'élément principal est le support de fixation des rosettes. L'axe optique du système doit passer par le centre du support des rosettes. L'ajustement du support des rosettes se fait donc en tout premier lieu. Il faut orienter le faisceau laser non collimé de façon à ce que celui-ci passe par le centre du support tout en étant perpendiculaire à sa surface avant. Il faut ensuite ajuster la distance entre la dernière lentille et le support des rosettes. C'est de cette distance que dépend l'espacement entre les répliques de l'image d'entrée. Pour positionner la dernière lentille à la bonne distance du support il faut mettre en place le réseau de diffraction. La distance adéquate est celle qui fait en sorte que les répliques du faisceau laser interceptent la surface avant du support approximativement aux positions prévues pour les centres des rosettes.

Une fois la dernière lentille placée à une distance convenable du support, les autres composants peuvent être mis en place et alignés selon l'axe optique du système (faisceau laser). La distance entre l'objectif téléphoto et la dernière lentille est approximative, toutefois la distance séparant les deux lentilles du téléphoto doit être ajustée de manière à obtenir un facteur de grandissement d'environ 1.05 X.

Le positionnement des rosettes se fait à l'aide d'un microscope et d'une mire placée au centre du plan d'entrée du montage (plan du ACL). Le microscope est doté d'une caméra vidéo pour faciliter l'ajustement. À tour de rôle, chacune des rosettes est centrée et orientée sur l'image de la mire à l'aide du mécanisme de positionnement. Le positionnement des filtres spatiaux se fait lorsque l'ACL est en place. Cette opération se fait également avec le microscope.

Le montage final est illustré à la figure 3.10. Des gros plans de l'ACL, du réseau de diffraction et des plans des filtres spatiaux et des rosettes apparaissent également à la figure 3.10.

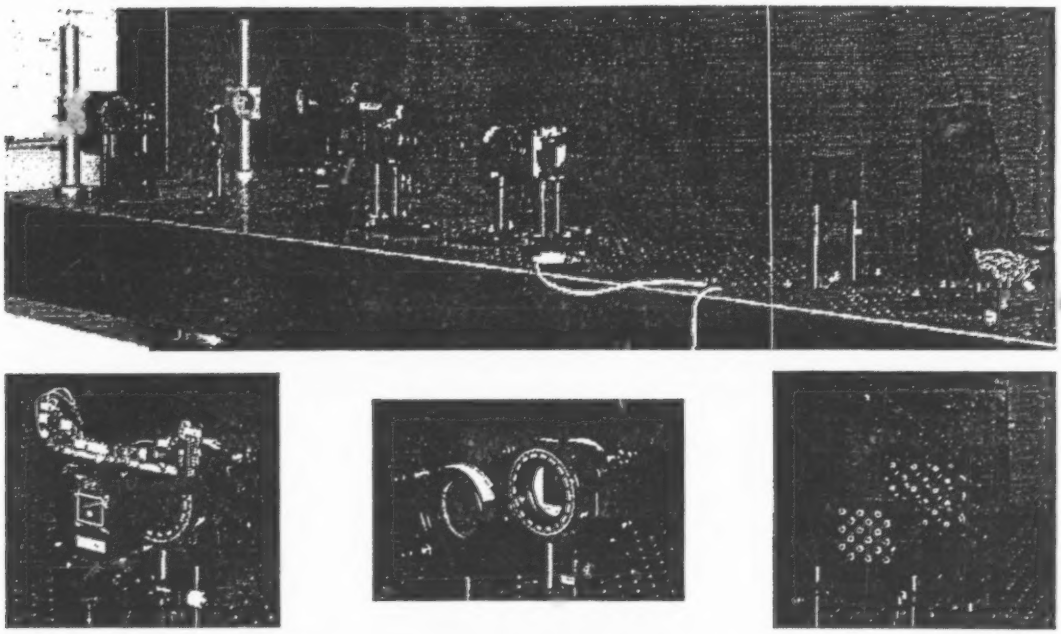


FIGURE 3.10 : Photographies du système complété avec gros plans en bas à droite sur l'ACL, en bas au centre sur le réseau de diffraction et en bas à gauche sur les filtres spatiaux et les rosettes.

3.2 Montage à deux ACL

Etant donné les problèmes inhérents au montage à réseau de diffraction, un deuxième montage pour l'extraction des CHC basé sur deux ACL a été considéré. L'objet segmenté est d'abord affiché sur le premier ACL. Celui-ci est éclairé par une source de lumière incohérente et est ensuite imagé sur le second ACL avec un grandissement unitaire. Les différentes rosettes sont affichées successivement sur le second ACL. L'énergie lumineuse transmise par les deux ACL est recueillie sur un détecteur par le biais d'une lentille collectrice. La figure 3.11 illustre le schéma du montage.

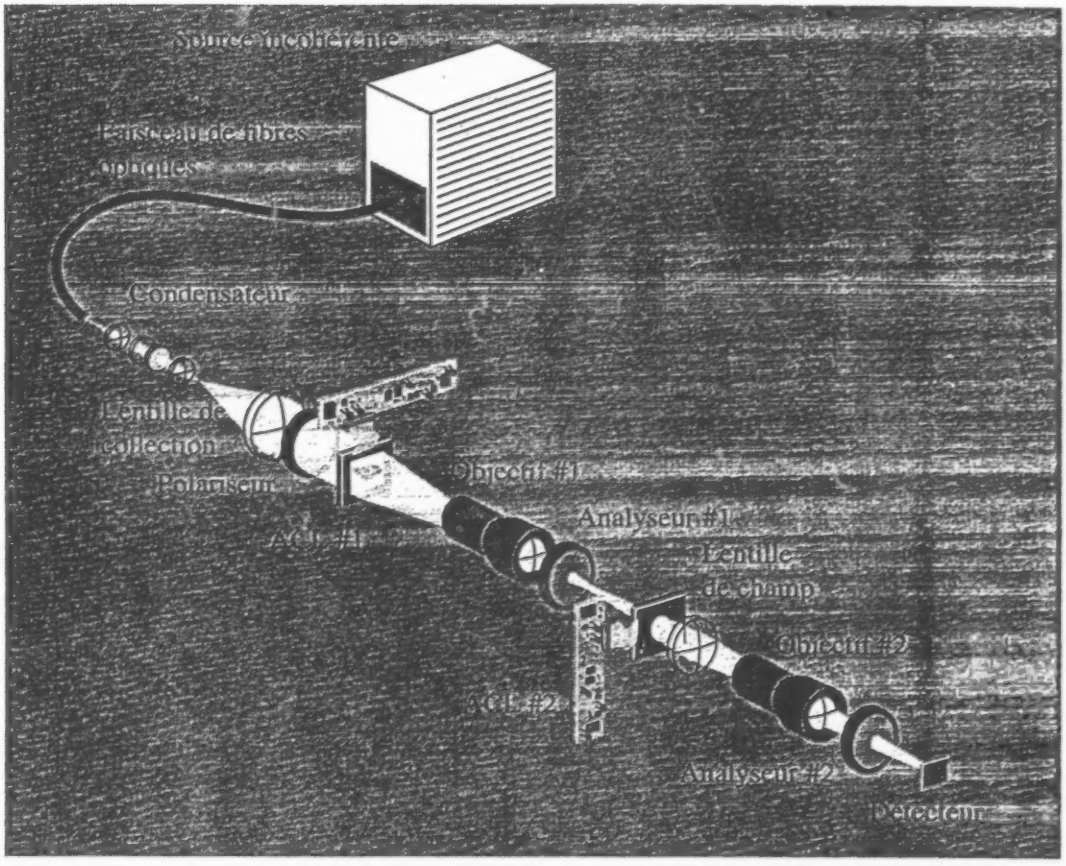


FIGURE 3.11 : Schéma du système à imagerie.

La source de lumière incohérente est constituée d'une ampoule incandescente et d'un faisceau de fibres optiques. L'utilisation d'un faisceau de fibres optique permet de réduire les inhomogénéités du faisceau émis par l'ampoule. Le système d'illumination comprend un condensateur (2 lentilles plano-convexes de 32 mm de focale) et d'une lentille de collection qui image la source dans le plan de la pupille d'entrée de l'objectif #1 (voir figure 3.11) . Pour améliorer l'uniformité de l'illumination un diffuseur holographique est intercalé entre les lentilles du condensateur. Un filtre d'absorption du rayonnement infrarouge à également été incorporé au condensateur pour protéger les ACL.

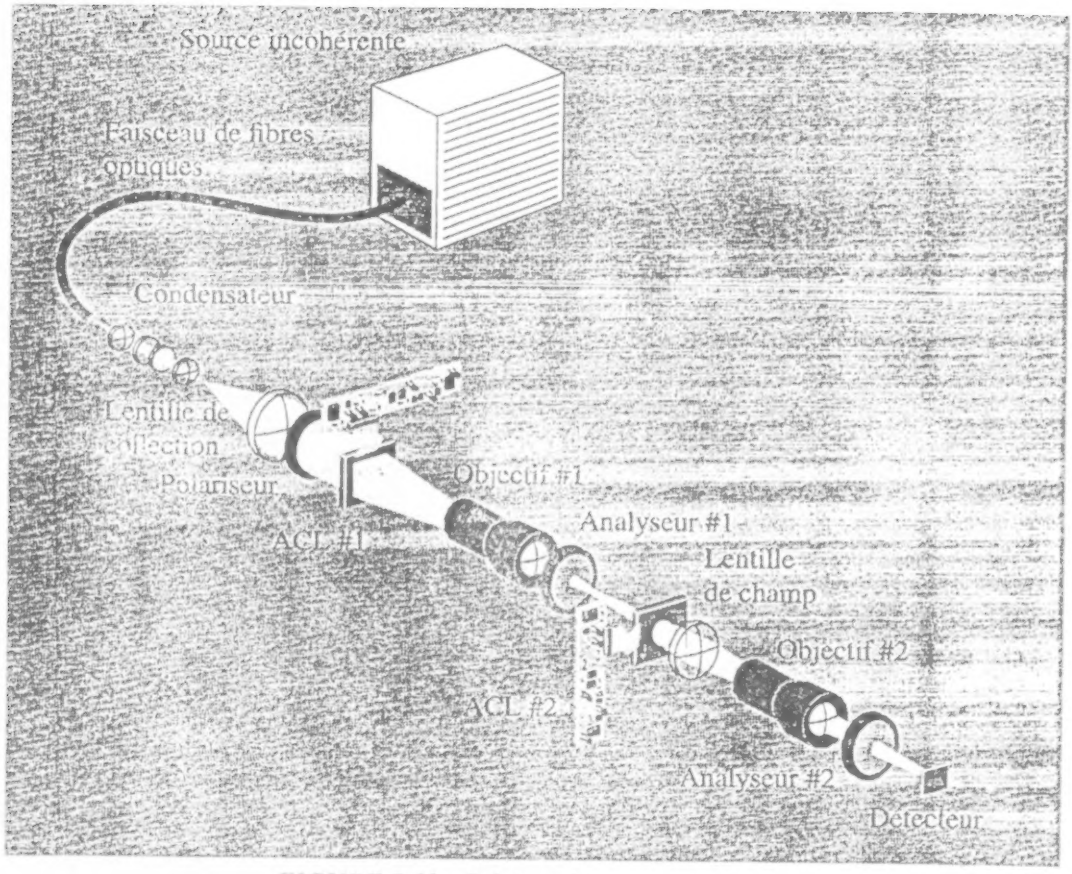


FIGURE 3.11 : Schéma du système à imagerie.

La source de lumière incohérente est constituée d'une ampoule incandescente et d'un faisceau de fibres optiques. L'utilisation d'un faisceau de fibres optique permet de réduire les inhomogénéités du faisceau émis par l'ampoule. Le système d'illumination comprend un condensateur (2 lentilles plano-convexes de 32 mm de focale) et d'une lentille de collection qui image la source dans le plan de la pupille d'entrée de l'objectif #1 (voir figure 3.11). Pour améliorer l'uniformité de l'illumination un diffuseur holographique est intercalé entre les lentilles du condensateur. Un filtre d'absorption du rayonnement infrarouge à également été incorporé au condensateur pour protéger les ACL.

L'objectif #1, qui image le premier ACL sur le second, est un *Apo-Rodagon D* de 75 mm de focale. Celui-ci est fabriqué par *Rodenstock*, il est conçu pour faire de l'imagerie avec un grandissement unitaire (1 X). C'est un objectif présentant une très bonne qualité optique et, fait très important, il produit des images pratiquement exemptes de distorsion. Le second objectif est identique au premier, toutefois un objectif de moindre qualité aurait suffit à la tâche puisque les spécifications requises sont beaucoup moins sévères. Le second objectif a pour fonction d'imager le second ACL sur le détecteur. Pour accroître l'énergie lumineuse collectée sur le détecteur, une lentille de champ est placée juste en arrière du second ACL. Cette lentille sert à imager la pupille du premier objectif, dans la pupille du second.

La lumière venant de la source est polarisée linéairement à l'aide d'un polariseur dichroïque placé entre la lentille de collection et le premier ACL. L'image affichée sur le premier ACL est polarisée grâce au premier analyseur. La direction de polarisation de ce dernier est approximativement à angle droit de celle du polariseur. À la sortie du premier analyseur, la lumière est polarisée linéairement mais sa direction de polarisation est perpendiculaire à celle de la lumière qui illumine le premier ACL. Pour cette raison, et de façon à éviter l'utilisation d'une composante optique de rotation de la polarisation[1], le second ACL est tourné de 90^0 relativement au premier ACL.

Puisqu'il y a une superposition de deux figures échantillonnées, des effets de Moiré très prononcés apparaissent si le premier ACL est imagé avec un grandissement non unitaire. Les effets de Moiré font également leur apparition lorsqu'il y a de la distorsion dans l'image. L'alignement des composants doit donc être très bien fait pour réduire les aberrations. De plus les distances entre les ACL et le premier objectif doivent être très bien ajustées, de façon à obtenir un grandissement rigoureusement unitaire.

Lorsque l'ajustement mécanique des ACL est complété, des mires similaires sont affichés simultanément sur les deux ACL. Le centre d'affichage de chaque ACL est déplacé électroniquement jusqu'à ce que les deux mires affichées coïncident avec une précision meilleure que 16μm.

1. Il en existe qui fournissent de très bons résultats pour des faisceaux monochromatiques bien collimés, toutefois pour un faisceau divergent polychromatique les résultats sont moins efficaces.

Une image de compensation est superposées aux rosettes, pour compenser le vignetage et la non uniformité de l'illumination. Cette image est calculée à partir de l'image du second ACL produit par le second objectif. Cette dernière image est capturée à l'aide d'une caméra vidéo placée temporairement approximativement à l'endroit où l'on retrouve normalement le photodétecteur. L'opération d'uniformisation est effectuée une seule fois lors de l'assemblage du montage. Le montage complété est illustré à la figure 3.12.

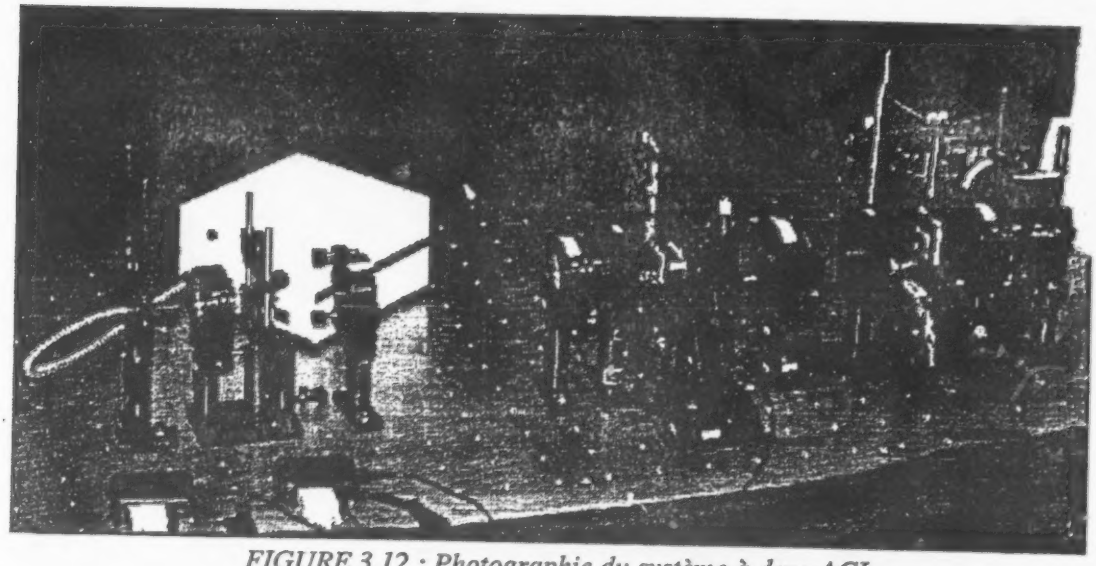


FIGURE 3.12 : Photographie du système à deux ACL.

3.3 Appréciation globale des deux montages optiques

3.3.1 Système à réseau de diffraction

Le système basé sur le réseau de diffraction est un système relativement complexe et délicat à réaliser. Ce système souffre d'un inconvénient important, celui de fonctionner en lumière cohérente. Les systèmes cohérents sont très sensibles à la présence de poussières déposées sur les composants. Les poussières dégradent significativement la qualité des

répliques de l'image d'entrée. Les images de la figure 3.13 illustre la dégradation causée par les poussières déposées sur les composants optiques du système à réseau de diffraction. L'image de gauche montre l'une des répliques produite par le système lorsque l'image d'entrée est un fond blanc uniforme. On constate la présence de nombreuses cannelures produites par la diffraction sur les granules de poussière. L'image à droite correspond à l'une des répliques lorsqu'est affichée à l'entrée l'image d'un modèle réduit.

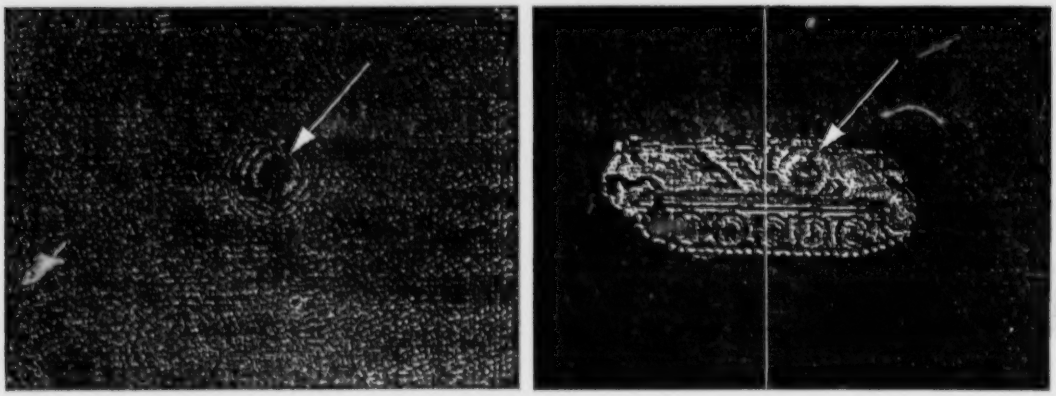


FIGURE 3.13 : Images prises à la sortie du système à réseau de diffraction avec à l'entrée un fond blanc uniforme en haut et une image de modèle réduit en bas à droite. La fleche montre une tache de diffraction qui affecte grandement l'apparance de l'image de l'objet.

Le système à réseau de.diffraction présente un autre désavantage: sa taille. De façon à limiter les aberrations optiques, des lentilles de focales relativement grandes ont dues être utilisées. De cet impératif il résulte un système mesurant plus de trois mètres de long. Le système pourrait être plié et une conception plus élaborée pourrait éventuellement permettre des réductions de volume significatifs. Toutefois il serait difficile de créer un système vraiment compacte. La faiblesse des signaux optiques qui parviennent aux détecteurs est un autre désavantage du système basé sur le réseau de diffraction. L'utilisation d'un laser plus puissant et de détecteurs plus sensibles permettrait d'améliorer le rapport signal à bruit du système. Toutefois cela rendrait le système encore plus complexe.

Il n'a pas été possible d'obtenir des résultats valables avec le système à réseau de diffraction. Une erreur s'est glissée lors de la création des fichiers informatiques utilisés pour la génération des rosettes. Le résultat de cette erreur est l'enregistrement de rosettes en $(1+cos)^{1/2}$ plutôt qu'en $1+cos$, ce qui s'est traduit évidement par une dégradation très significative des résultats. Il ne restait plus assez de temps pour réaliser d'autres rosettes sur masque en chrome. De toute façon il semble que le système à réseau de diffraction ne soit pas bien adapté au calcul des caractéristiques CHC. Le calcul des CHC requiert des mesures relativement précises des valeurs des rosettes. Le bruit associé à l'éclairement cohérent limite grandement la précision des mesures. Pour obtenir des résultats valables avec le système à réseau de diffraction il faudrait consacré des efforts important pour protéger le système des poussières. Il faudrait également faire une conception optique élaborée spécifique pour le système.

3.3.2 Montage à deux ACL

Le montage à deux ACL malgré ses défauts, à permis d'obtenir des résultats intéressants. Le principal défaut du montage utilisé réside dans la non synchronisation des ACL. L'équipement utilisé ne permettait pas l'adressage des ACL de façon synchrone. L'affichage non simultané de l'objet d'entrée et des rosettes se traduit par un phénomène de battement qui fausse les mesures des valeurs des rosettes. La correction du problème de la synchronisation amenerait des améliorations significatives sur la précision des mesures.

Dans sa configuration actuelle le système permet l'évaluation d'une seule valeur de rosette par cycle vidéo. Compte tenus des mesures supplémentaires requises pour compenser certaines erreurs, l'évaluation d'une signature à 7 caractéristiques CHC nécessiteraient plus d'une demi seconde. Pour en arriver à un traitement en un seul cycle vidéo de 1/30 sec, il faudrait prévoir plusieurs ACL affichant chacun plusieurs répliques de l'image d'entrée. Puisqu'une résolution de 200x200 pixels serait suffisante et que les ACL disposent de 640x480 pixels, il serait possible d'évaluer au moins quatre valeurs de rosette par ACL (soit deux CHC). Ainsi l'évaluation de sept caractéristiques CHC nécessiterait l'utilisation de quatre montages.

Le système à deux ACL présente une assez bonne qualité au niveau de l'optique. L'obtention d'images de qualité est grandement simplifiée par le fait que le système opère avec une illumination incohérente. Le système à deux ACL est relativement simple et compacte, toutefois ses fonctions pourraient très probablement être reproduites dans un système plus simple et beaucoup plus compacte. La figure 3.14 montre un tel système très compacte. Ce système comprendrait un seul ACL sur lequel serait collé des rosettes de hautre résolution. Un tel système serait guère plus volumineux que l'ACL lui-même. On pourrait éventuellement mettre plusieurs de ces systèmes sur une carte de circuit imprimé de format carte PC.

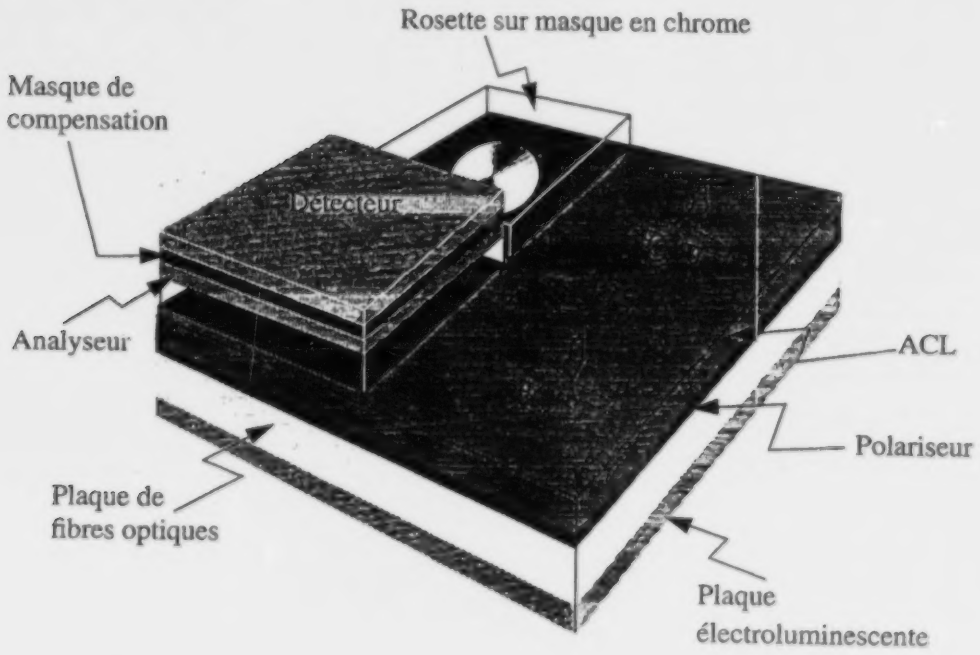


FIGURE 3.14 : Système compact envisageable d'extraction des caractéristiques CHC et CHCA.

En conclusion on peut dire que l'approche illustrée par le système à deux ACL est celle qui offre les meilleurs perspectives.

CHAPITRE 4

Traitement numérique

Ce chapitre présente les travaux reliés au traitement numérique des images vidéo. Il se divise en deux sections. La première section expose la préparation des images en vue de l'extraction des CHC, alors que la seconde aborde le post-traitement.

4.1 Préparation des images

Cette section décrit le traitement numérique des images vidéo précedent l'extraction des CHC par le système optique. La préparation des images a été simplifiée de façon importante suite à la décision de limiter l'étude à des images vidéo provenant d'une scène contrôlée en laboratoire. De façon générale, le traitement numérique développé ici se résume aux étapes énoncées à la figure 4.1.

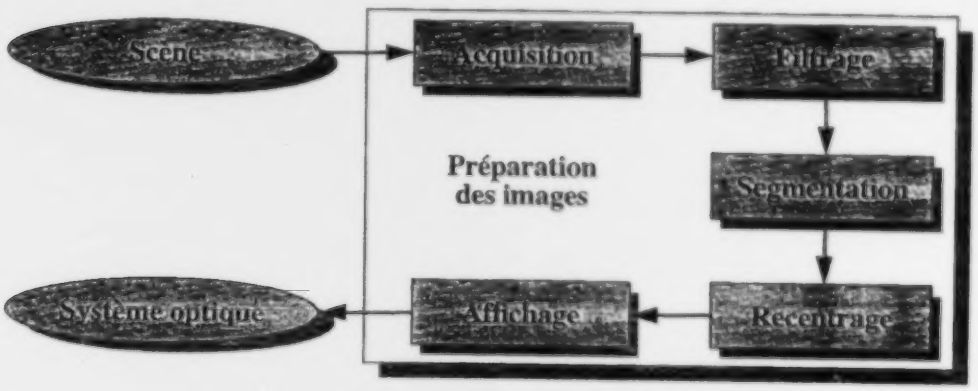


FIGURE 4.1 : Les étapes de la préparation des images.

Le système à deux ACL présente une assez bonne qualité au niveau de l'optique. L'obtention d'images de qualité est grandement simplifiée par le fait que le système opère avec une illumination incohérente. Le système à deux ACL est relativement simple et compacte, toutefois ses fonctions pourraient très probablement être reproduites dans un système plus simple et beaucoup plus compacte. La figure 3.14 montre un tel système très compacte. Ce système comprendrait un seul ACL sur lequel serait collé des rosettes de hautre résolution. Un tel système serait guère plus volumineux que l'ACL lui-même. On pourrait éventuellement mettre plusieurs de ces systèmes sur une carte de circuit imprimé de format carte PC.

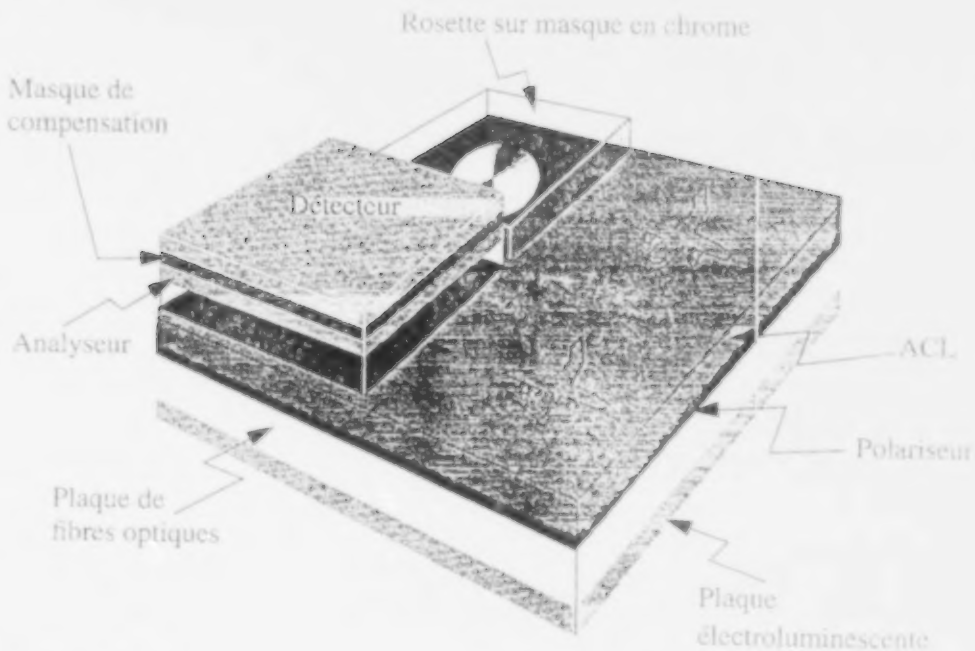


FIGURE 3.14 : Système compact envisageable d'extraction des caractéristiques CHC et CHCA.

En conclusion on peut dire que l'approche illustrée par le système à deux ACL est celle qui offre les meilleurs perspectives.

CHAPITRE 4

Traitement numérique

Ce chapitre présente les travaux reliés au traitement numérique des images vidéo. Il se divise en deux sections. La première section expose la préparation des images en vue de l'extraction des CHC, alors que la seconde aborde le post-traitement.

4.1 Préparation des images

Cette section décrit le traitement numérique des images vidéo précédent l'extraction des CHC par le système optique. La préparation des images a été simplifiée de façon importante suite à la décision de limiter l'étude à des images vidéo provenant d'une scène contrôlée en laboratoire. De façon générale, le traitement numérique développé ici se résume aux étapes énoncées à la figure 4.1.

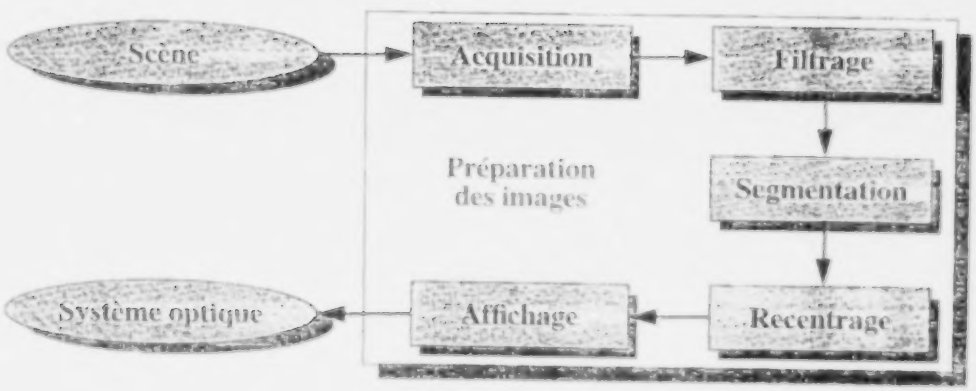


FIGURE 4.1 : Les étapes de la préparation des images.

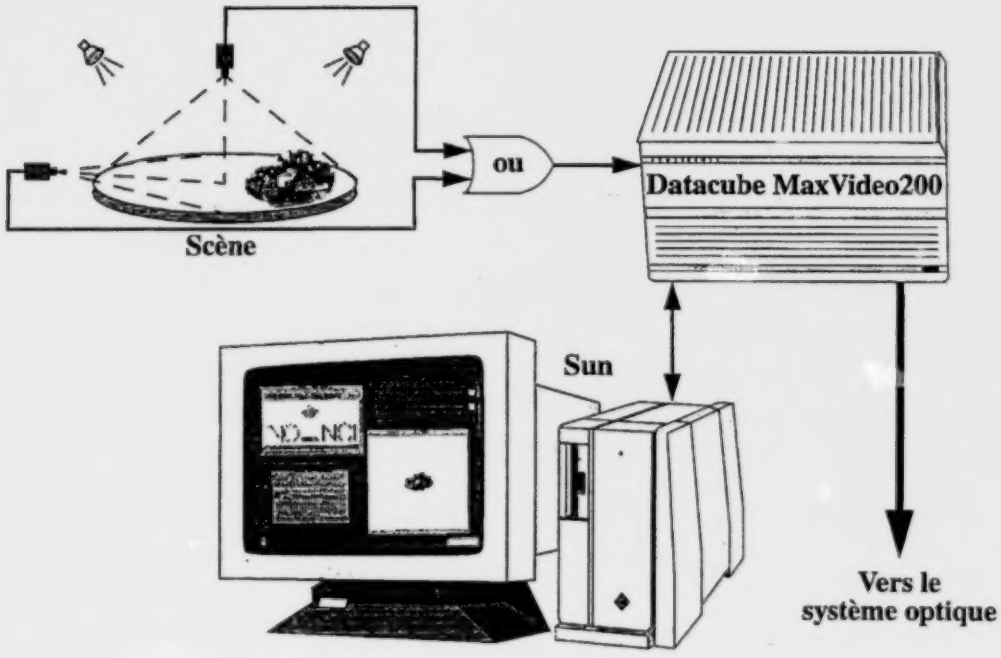


FIGURE 4.2 : Les composants de la préparation des images.

Chacune de ces opérations est exécutée sur l'architecture Datacube conjointement avec un ordinateur de type Sun SPARC 20 (voir figure 4.2). Le système Datacube que l'INO possède est muni d'une carte MaxVideo200 qui comprend les modules suivants[1]:

- "Analog Scanner" (*AS*)
- "Architectural Adapter version B" (*AB*)
- "Advanced Memory" (*AM*)
- "Analog Generator" (*AG*)
- "Advanced Pipeline Processor" (*AP*)
- "Acquire Digital" (*AD*)
- "Arithmetic Unit" (*AU*)

1. Seuls les quatre premiers modules sont utilisés (*AS*, *AB*, *AM* et *AG*).

Le traitement numérique que nous avons mis en oeuvre requiert l'utilisation de quatre pipelines distincts sur l'architecture Datacube. La figure 4.3 illustre ces pipelines avec leur élément d'origine et leur(s) élément(s) de destination. Les éléments d'information

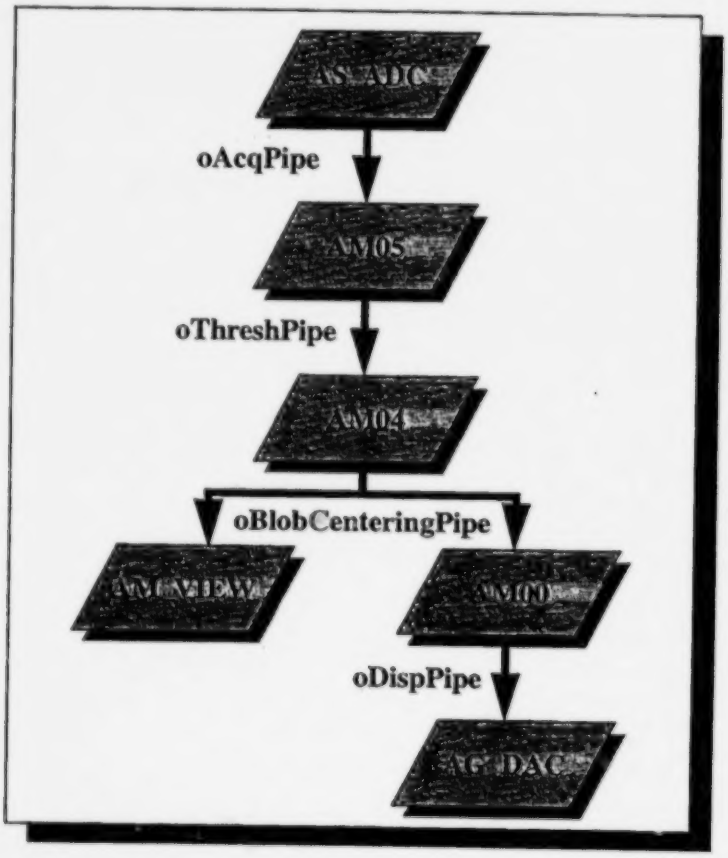


FIGURE 4.3 : Les pipelines de l'architecture de traitement.

supplémentaires sur leurs fonctions sont présentés dans les sections suivantes.

4.1.1 La scène

La scène utilisée est entièrement contrôlée. Nous en maîtrisons donc les principaux éléments dont la cible, l'arrière-scène et l'éclairage (voir figure 4.4). La cible est formée d'un seul modèle réduit de véhicule déposé sur un plateau rotatif qui tourne à vitesse constante. Le plateau rotatif se confond parfaitement à l'arrière-scène qui est d'un blanc uniforme. Le

Le traitement numérique que nous avons mis en oeuvre requiert l'utilisation de quatre pipelines distincts sur l'architecture Datacube. La figure 4.3 illustre ces pipelines avec leur élément d'origine et leur(s) élément(s) de destination. Les éléments d'information

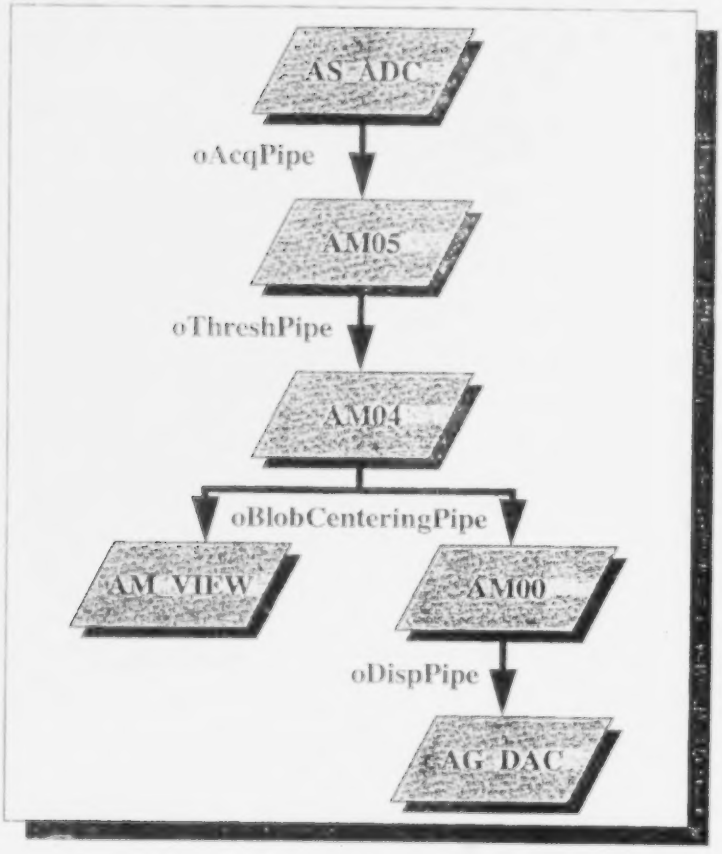


FIGURE 4.3 : Les pipelines de l'architecture de traitement.

supplémentaires sur leurs fonctions sont présentés dans les sections suivantes.

4.1.1 La scène

La scène utilisée est entièrement contrôlée. Nous en maîtrisons donc les principaux éléments dont la cible, l'arrière-scène et l'éclairage (voir figure 4.4). La cible est formée d'un seul modèle réduit de véhicule déposé sur un plateau rotatif qui tourne à vitesse constante. Le plateau rotatif se confond parfaitement à l'arrière-scène qui est d'un blanc uniforme. Le

système d'éclairage permet de faire ressortir clairement la cible tout en éliminant le plus possible les ombres aux abords de cette dernière.

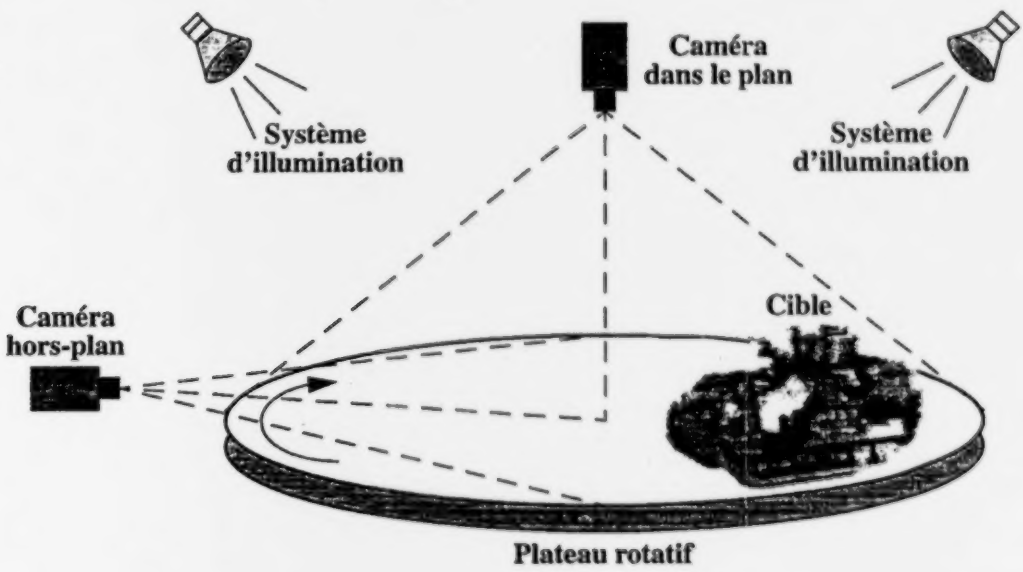


FIGURE 4.4 : Scène utilisée pour la préparation des images.

4.1.2 L'acquisition

L'acquisition a pour but de saisir et de numériser les images vidéo provenant de la scène. La saisie se fait à partir d'une caméra vidéo commerciale (Panasonic modèle GP-MF552). Cette caméra noir et blanc est munie d'un CCD de 2/3" et possède un rapport signal à bruit de 59 dB. La caméra est située au dessus du plateau rotatif, ou encore sur le côté, permettant ainsi d'obtenir respectivement des images vidéo dans le plan et hors-plan. L'une ou l'autre des configurations de caméra est utilisée, selon la prise de vue souhaitée.

La caméra transmet un signal vidéo de type RS-170 au système de traitement vidéo Datacube. Le module *AS* du Datacube numérise alors le signal à cadence vidéo (30 images par seconde) avec une résolution spatiale de 512 x 484 pixels; 8 bit/pixel (256 niveaux de gris).

Le pipeline d'acquisition (*oAcqPipe*) débute avec l'élément *AS_ADC* du module *AS* et se termine avec l'élément *AM_MEM_R8* du module *AM05*, en passant par le module *AB*. Ce pipeline fonctionne à cadence vidéo.

4.1.3 Filtrage

Le filtrage consiste à traiter les images acquises de façon à faciliter l'extraction ou la segmentation de la cible par rapport à l'arrière-scène. Dans le rapport de la première partie du projet, un filtre pseudo-médian avait été proposé en guise d'opérateur permettant de réaliser ce filtrage. Il avait été alors mentionné que ce filtre permettait de réduire le bruit non-corrélé souvent associé aux images infrarouges. Bien qu'intéressante en théorie, cette approche n'a toutefois pas été mise en oeuvre pour des raisons pratiques.

En effet, le filtre pseudo-médian est en réalité une approximation imparfaite du filtre médian dont la mise en oeuvre s'avère très difficile sur une architecture qui se veut temps-réel. Bien que plus simple à réaliser, la mise en oeuvre du filtre pseudo-médian sur l'architecture Datacube nécessite l'utilisation d'éléments particuliers appelés *MINMAX* qui ne se retrouvent malheureusement pas sur les modules Datacube que l'INO possède. La mise en oeuvre du filtre pseudo-médian aurait nécessité l'achat d'un nouveau module très dispendieux. Compte tenu des simplifications apportées à la scène, l'INO n'a pas cru bon de procéder à cet achat. En effet, la présence de bruit non-corrélé est fortement réduite dans une scène contrôlée, simplifiant ainsi l'opération de filtrage et éliminant la nécessité du filtrage pseudo-médian.

Dans le cadre du prétraitement numérique développé pour ce projet, le filtrage se limite donc à l'utilisation d'un filtre passe-bas réalisé à même le pipeline d'acquisition (*oAcqPipe*) sur l'élément *LPF* du module *AS* du système Datacube. Ce filtre fonctionne à taux vidéo et suffit à éliminer une bonne partie du bruit haute fréquence souvent lié à l'électronique d'acquisition. Il a de plus l'avantage de ne monopoliser qu'un seul élément du Datacube, ce qui évite l'ajout d'un pipeline dédié.

4.1.4 Segmentation

La segmentation est l'opération de traitement qui permet d'isoler la cible par rapport à l'arrière-scène. Cette étape est aussi grandement simplifiée par le fait que nous avons limité notre étude à une scène contrôlée. Puisque l'arrière-scène est composée d'un blanc uniforme qui sature la caméra, la segmentation se résume à seuiller les images de façon à conserver uniquement les pixels qui possèdent une intensité inférieure à celle de l'arrière-scène, soit 255. Notre connaissance *a priori* de la scène nous permet d'estimer que ces pixels appartiennent à la cible. L'opération de seuillage est décrite à la figure 4.5.

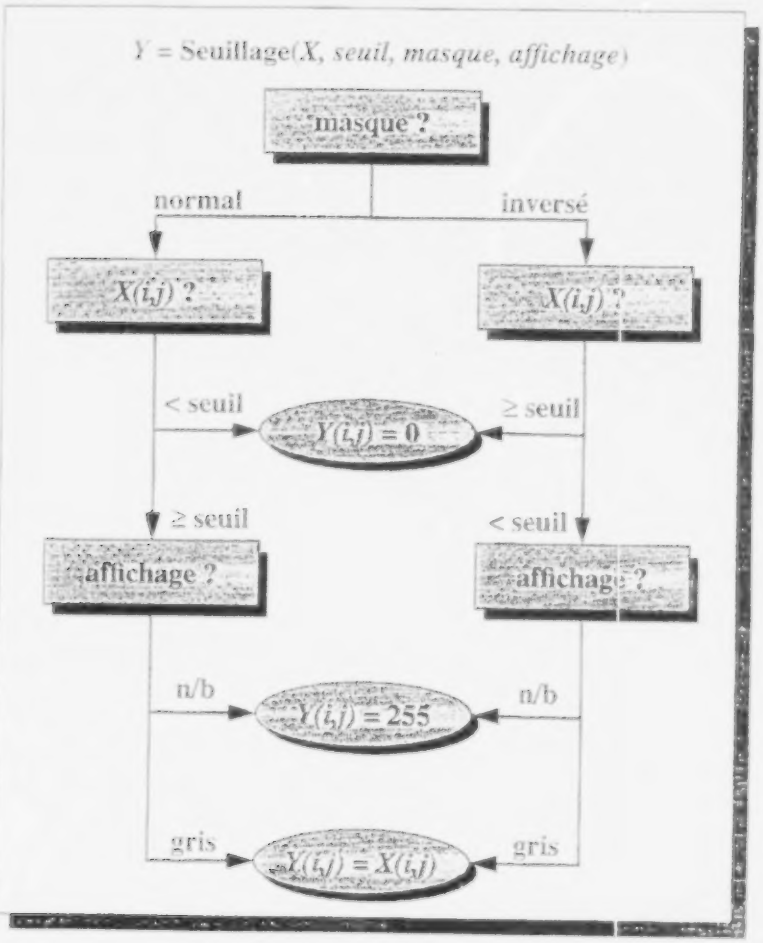


FIGURE 4.5 : Schéma de l'opération de seuillage.

4.1.4 Segmentation

La segmentation est l'opération de traitement qui permet d'isoler la cible par rapport à l'arrière-scène. Cette étape est aussi grandement simplifiée par le fait que nous avons limité notre étude à une scène contrôlée. Puisque l'arrière-scène est composée d'un blanc uniforme qui sature la caméra, la segmentation se résume à seuiller les images de façon à conserver uniquement les pixels qui possèdent une intensité inférieure à celle de l'arrière-scène, soit 255. Notre connaissance *a priori* de la scène nous permet d'estimer que ces pixels appartiennent à la cible. L'opération de seuillage est décrite à la figure 4.5.

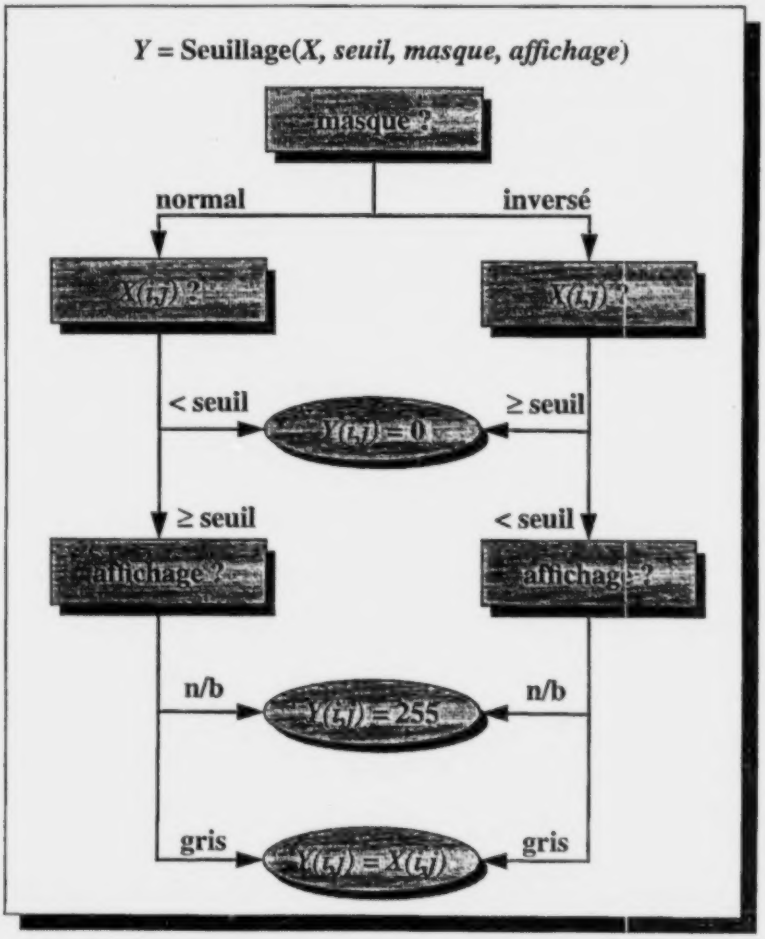


FIGURE 4.5 : Schéma de l'opération de seuillage.

En plus du *seuil*, l'opération de seuillage possède deux autres paramètres qui permettent de faire varier les résultats de la segmentation: le *masque* et l'*affichage*. D'abord, lorsque le paramètre *masque* est *normal*, tous les pixels (i,j) de l'image d'entrée (X) qui possèdent une intensité inférieure au *seuil* sont mis à zéro dans l'image de sortie (Y). De même, lorsque le *masque* est *inversé*, tous les pixels (i,j) de l'image d'entrée (X) qui possèdent une intensité supérieure ou égale au *seuil* sont mis à zéro dans l'image de sortie (Y). Selon le type d'*affichage* choisi, les autres pixels de l'image de sortie sont alors mis à 255 (*affichage* en *noir et blanc*) ou demeure inchangés (*affichage* en *niveaux de gris*).

En général, on désire segmenter la cible de façon à afficher cette dernière en niveaux de gris sur fond noir. Cette segmentation est réalisée en appliquant par défaut le seuillage suivant:

$$Y = \text{Seuillage}(X, 255, \text{inversé}, \text{gris}). \tag{4.1}$$

La mise en oeuvre de l'opération de seuillage sur le système Datacube se fait au moyen de l'élément *LUT* situé sur le module *AB*. Le pipeline de segmentation (*oThreshPipe*) formé débute à l'élément *AM_MEM_R8* du module *AM05*, passe à travers l'élément *LUT* du module *AB*, et se termine à l'élément *AM_MEM_R8* du module *AM04* (voir figure 4.3). Le pipeline de segmentation fonctionne à cadence vidéo.

4.1.5 Recentrage

Le recentrage est l'opération qui permet de repositionner la cible au centre de l'image. Cette étape est rendue nécessaire en raison du fait que le système optique utilise directement les images plutôt que la transformée de Fourier de ces dernières.

Le recentrage consiste en premier lieu à déterminer la position du centre géométrique de la silhouette de la cible à partir de l'image issue de la segmentation. Le centre géométrique de la cible (I_g, J_g) est obtenu en calculant les indices i et j moyens des N pixels appartenant à la cible. Cela s'écrit:

$$\forall (i,j) \in \text{cible}, \quad I_g = \frac{1}{N}\sum i \quad \text{et} \quad J_g = \frac{1}{N}\sum j. \tag{4.1}$$

Si l'image est constituée d'une seule cible, l'opération de recentrage peut être presque entièrement réalisée sur le système Datacube en utilisant l'élément *AM_VIEW* situé sur l'un des modules *AM* de la carte MaxVideo200. Cet élément permet de calculer à taux vidéo les sommes en i et j, de même que le nombre total de pixels compris dans la cible (N). Le Sun SPARC20 est utilisé pour compléter cette opération en effectuant les divisions par N nécessaires au calcul du centre de gravité. L'utilisation du Sun SPARC20 ne ralentit pas pour autant le traitement puisque le transfert des trois valeurs entières (les deux sommations et N) du Datacube au Sun SPARC20 ainsi que les deux divisions par N qui suivent peuvent aisément être réalisés à cadence vidéo.

En second lieu, une fois le centre géométrique (I_g, J_g) de la cible obtenu, il faut déplacer la cible vers le centre de l'image (I_c, J_c). Cette opération (voir figure 4.6) se résume à décaler l'image entière d'un certain vecteur de translation $\vec{T}$ donné par:

$$\vec{T} = [(I_c - I_g), (J_c - J_g)] . \tag{4.1}$$

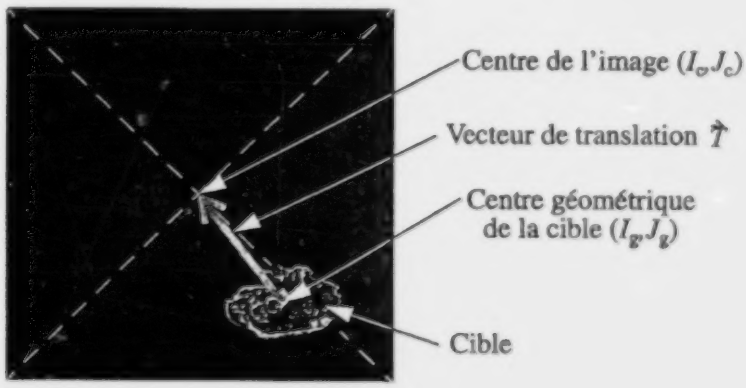


FIGURE 4.6 : Recentrage par translation.

En pratique cette translation consiste simplement à modifier le point d'alignement de la surface ("surface alignment point") de l'élément *AM_MEM_R8* du module *AM00* où est situé l'image avant l'affichage. Cette modification s'effectue d'une image à l'autre dès le centre de gravité de la cible calculé.

Le pipeline conçu pour effectuer le recentrage possède deux destinations ("multidestination pipe") et s'appelle *oBlobCenteringPipe*. Il part de l'élément *AM_MEM_R8* du module *AM04*, passe à travers le module *AB*, et se termine aux éléments *AM_VIEW* et *AM_MEM_R8* du module *AM00* (voir figure 4.3).

4.1.6 Affichage

L'affichage est la dernière étape de la préparation des images. Cette étape consiste à transmettre correctement les images au système optique et plus particulièrement à l'afficheur à cristaux liquides.

L'afficheur à cristaux liquides utilisé accepte comme entrée un signal de type "VGA non-interlaced". Le module *AG* de la carte MaxVideo200 du Datacube permet de générer facilement un tel signal. Pour ce faire, l'élément *AG_DAC* du module *AG* doit être utilisé en mode 512 x 512 pixels, 8 bit/pixel et "non-interlaced".

Le pipeline d'affichage (*oDispPipe*) débute à l'élément *AM_MEM_R8* du module *AM00* et se termine à l'élément *AG_DAC* du module *AG* après avoir circulé à travers le module *AB*, le tout se faisant à taux vidéo (voir figure 4.3).

4.1.7 Environnement de travail

Le système Datacube est contrôlé par le Sun SPARC20 via un interface usager que nous avons développé sur mesure (figure 4.7). Cet interface usager permet d'ajuster les différents paramètres du traitement, tel que décrit précédemment.

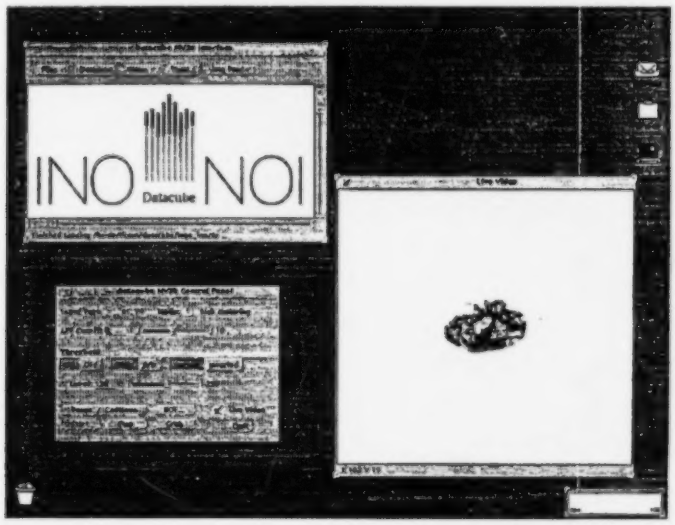


FIGURE 4.7 : Environnement de travail de la préparation des images.

4.1.8 Résultats de la préparation des images

La figure 4.8 illustre les résultats obtenus au cours des différentes étapes de la préparation des images.

En haut à gauche de la figure 4.8, on retrouve l'image originale, obtenue immédiatement après l'acquistion et la numérisation. On y observe un peu de bruit engendré par la caméra et les différentes composantes électroniques du circuit d'acquisition.

En haut à droite de la figure 4.8, on observe l'image filtrée issue du filtre passe-bas situé à même le circuit d'acquisition. On y remarque une réduction appréciable de l'information à hautes fréquences spatiales.

En bas à droite de la figure 4.8, apparait le résultat de la segmentation. On y observe l'élimination non désirée de certaines petites parties de l'objet. Ces parties correspondent à des régions de l'objet qui sont saturées par le système d'illumination et donc éliminées par le seuillage au même titre que l'arrière-scène.

Finalement, en bas à gauche de la figure 4.8, on retrouve l'image finale issue du recentrage de l'objet segmenté. C'est cette image qui est présentée au système optique d'extraction des CHC.

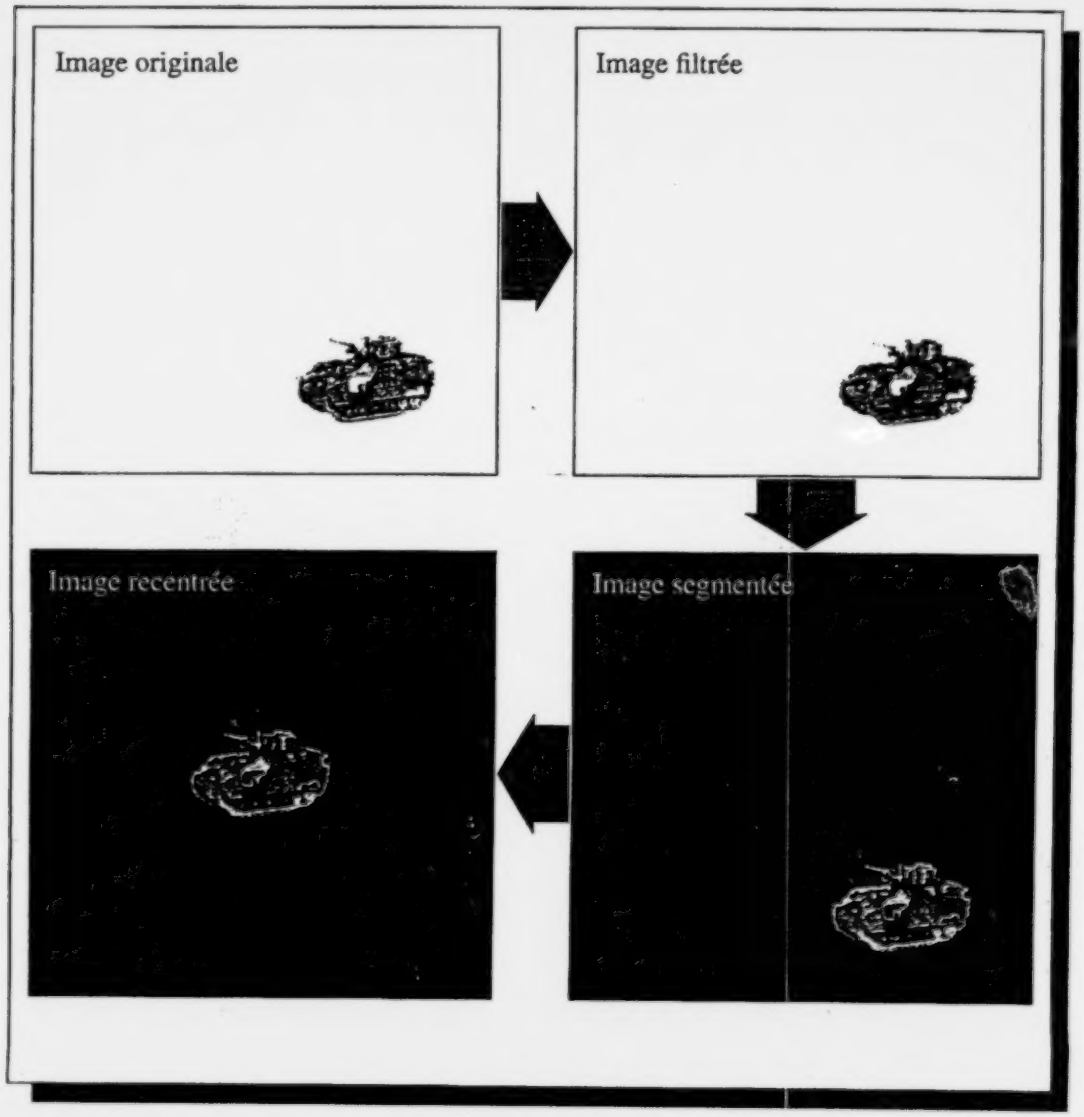


FIGURE 4.8 : Images obtenues à différentes étapes de la préparation des images.

4.2 Post-traitement

Cette section décrit les étapes du traitement numérique menant de l'acquisition des données à la classification de la cible, suite d'opérations appellé post-traitement. Le post-traitement se divise en quatre étapes, telles que montrées à la figure 4.9.

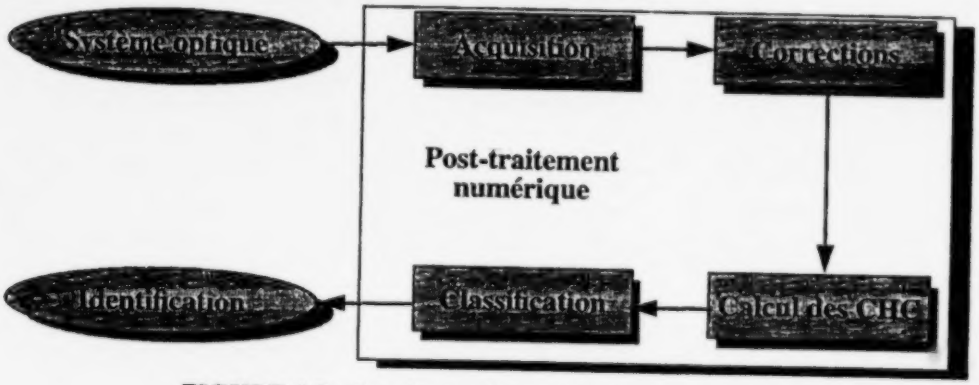


FIGURE 4.9 : Les étapes du post-traitement numérique.

4.2.1 Acquisition des données

Les données sont capturées sur un ordinateur de type PC, à l'aide d'une carte d'acquisition AT-MIO-64E3 de *National Instruments*. Cette carte dispose de 64 canaux analogiques de lecture/écriture, entièrement reconfigurables par commande logicielle. En mode différentiel, deux canaux sont utilisés pour chaque détecteur afin de réduire les bruits électroniques.

La carte admet des tensions analogiques sur la plage de 0-10V, correspondant aux signaux lus aux détecteurs, puis passés par des étages de pré-amplification et d'amplification. La conversion numérique se fait sur 12 bits. Les cadences maximales de lecture sont de 500K échantillons/s sur canal simple, et de 300K échantillons/s sur plusieurs canaux. Ces cadences élevées permettent de réaliser plusieurs lectures sur chaque canal (détecteur) à l'intérieur de la période d'une trame vidéo (33ms).

Nous avons réalisé un logiciel de commande et d'interface usager sous Windows 3.1 pour cette carte d'acquisition. Il permet de reconfigurer à volonté les paramètres d'acquisition

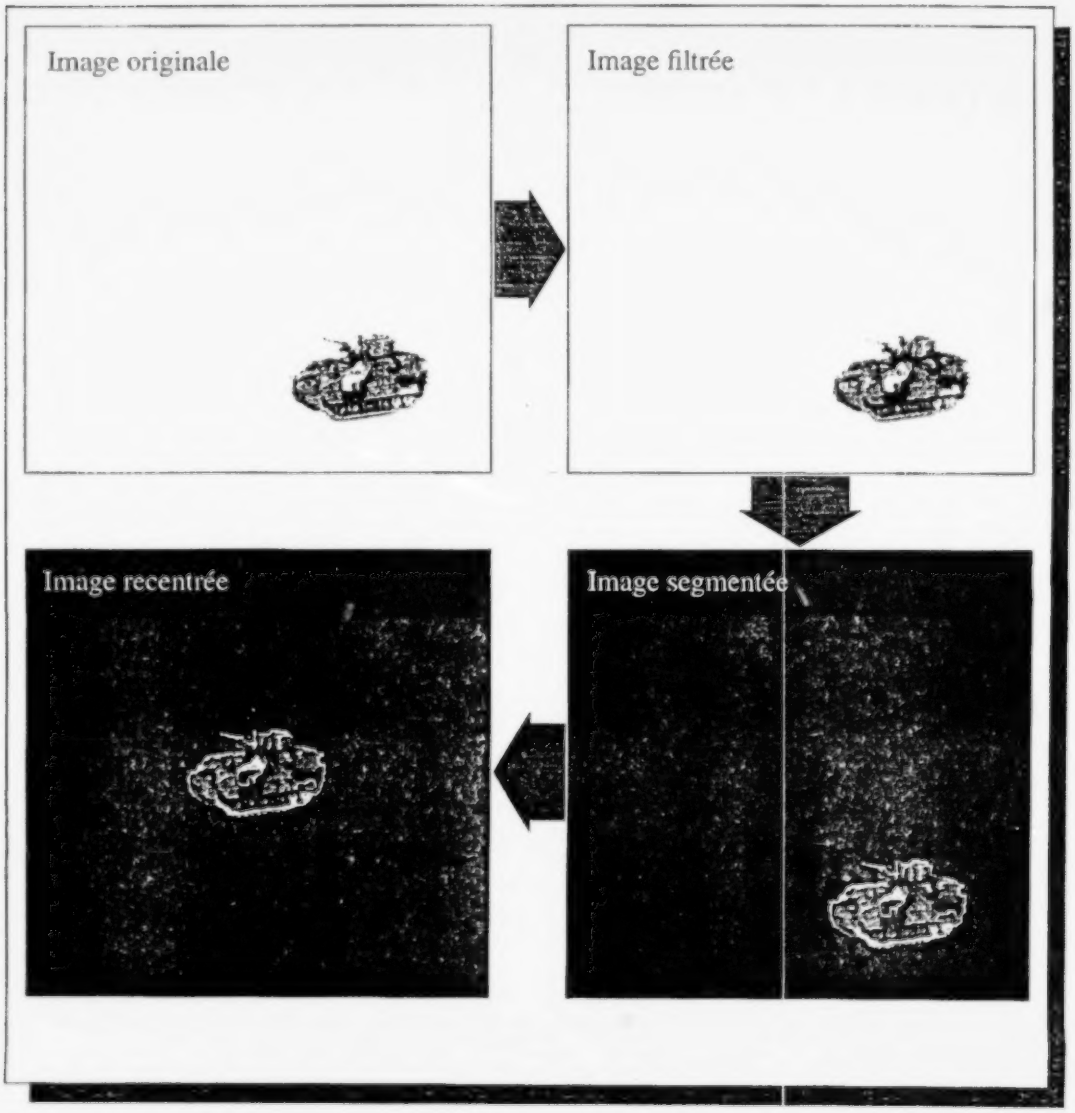


FIGURE 4.8 : Images obtenues à différentes étapes de la préparation des images.

4.2 Post-traitement

Cette section décrit les étapes du traitement numérique menant de l'acquisition des données à la classification de la cible, suite d'opérations appellé post-traitement. Le post-traitement se divise en quatre étapes, telles que montrées à la figure 4.9.

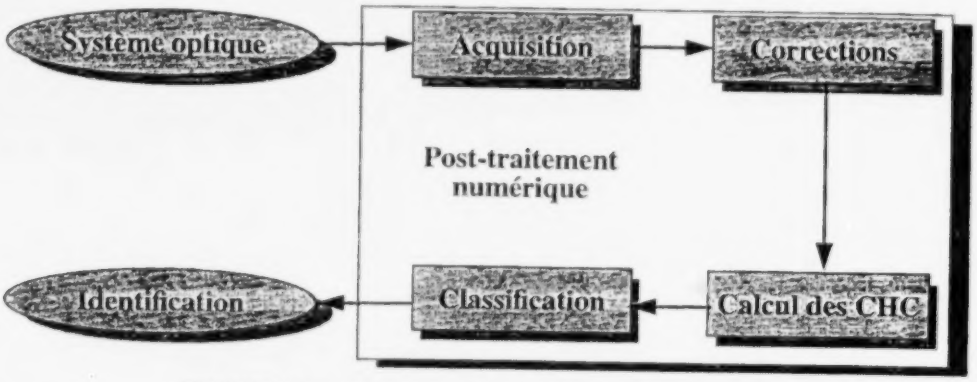


FIGURE 4.9 : Les étapes du post-traitement numérique.

4.2.1 Acquisition des données

Les données sont capturées sur un ordinateur de type PC, à l'aide d'une carte d'acquisition AT-MIO-64E3 de *National Instruments*. Cette carte dispose de 64 canaux analogiques de lecture/écriture, entièrement reconfigurables par commande logicielle. En mode différentiel, deux canaux sont utilisés pour chaque détecteur afin de réduire les bruits électroniques.

La carte admet des tensions analogiques sur la plage de 0-10V, correspondant aux signaux lus aux détecteurs, puis passés par des étages de pré-amplification et d'amplification. La conversion numérique se fait sur 12 bits. Les cadences maximales de lecture sont de 500K échantillons/s sur canal simple, et de 300K échantillons/s sur plusieurs canaux. Ces cadences élevées permettent de réaliser plusieurs lectures sur chaque canal (détecteur) à l'intérieur de la période d'une trame vidéo (33ms).

Nous avons réalisé un logiciel de commande et d'interface usager sous Windows 3.1 pour cette carte d'acquisition. Il permet de reconfigurer à volonté les paramètres d'acquisition

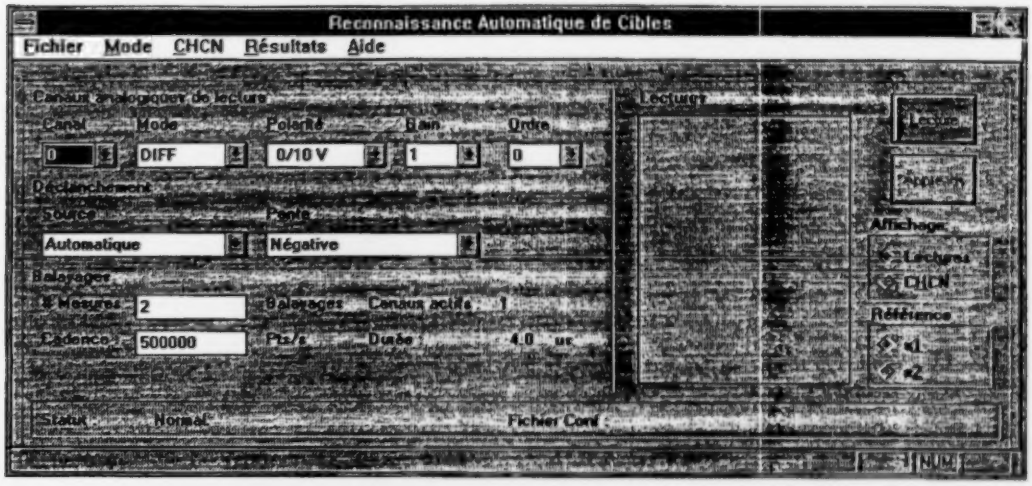


FIGURE 4.10 : Interface usager du logiciel de commande de la carte d'acquisition.

et d'effectuer la prise des mesures. La fenêtre principale du logiciel est présentée à la figure 4.10.

Les options du menu "Fichier", donnent accès aux commandes de chargement et de sauvegarde d'une configuration de la carte, ainsi qu'à la sortie du logiciel.

Le menu "Mode" contient les modes d'opération manuel et automatique. Cependant, les difficultés rencontrées avec les montages expérimentaux ont monopolisé nos efforts, ne nous permettant pas de réaliser les branchements nécessaires à la synchronisation des traitements. Le mode automatique n'est donc pas opérationel.

Le menu "CHCN" fait apparaître le dialogue présenté à la figure 4.11. Il permet d'associer les canaux de lecture correspondant aux filtres en cosinus et en sinus de chaque ordre, et de spécifier le canal identifié à l'ordre zéro. Cette association effectuée, le programme pourra faire automatiquement le calcul des CHC, puis afficher les résultats (voir la suite). Finalement, le menu "Résultats" permet de sauvegarder sur disque les lectures prises, dans un fichier nommé LECT.RES .

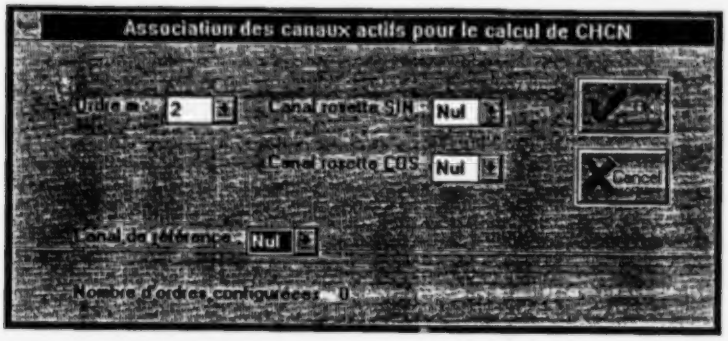


FIGURE 4.11 : Dialogue permettant l'association des canaux de lecture pour fin de calcul des CHC.

La configuration individuelle de chaque canal est réalisée à l'aide des cinq boîtes regroupées sous la rubrique "Canaux analogiques de lecture". La synchronisation des lectures avec le signal vidéo qui commande l'affichage de la scène dans le système optique peut-être programmée à l'aide des boîtes mises sous la rubrique "Déclanchement". Le nombre de mesures par canal, ainsi que la cadence de mesure sont régies par les valeurs entrées sous la rubrique "Balayage". On retrouve aussi sous cette rubrique les informations sur le nombre de canaux actifs et la durée totale des balayages.

Lorsque l'utilisateur appuie sur le bouton "Lecture", en haut à droite, l'ordre est donné de réaliser les balayages demandés sur les canaux actifs. Lorsque le bouton radio "Lectures" du groupe "Affichage" est activé, les mesures brutes sont récupérées, puis la moyenne et la variance dans chaque canal s'affiche dans la boîte marquée "Lectures". Lorsqu'au contraire c'est le boutons radio "CHCN" qui est activé, ce sont les valeurs moyennes des CHC calculés qui s'affichent dans la même boîte que précédemment, rebaptisée alors "CHCN".

Le second groupe de boutons radios apparaissant en bas à droite s'appelle "Référence". Il permet de pondérer le canal correspondant à l'ordre zéro par un facteur 1 ou 2, selon le besoin.

4.2.2 Corrections des données

Les mesures moyennées constituent la donnée χ_k pour le canal k.

Les données brutes doivent être corrigées pour tenir compte des variations de niveau et de sensibilité des différents canaux de détection. A cette fin, les mesures tirées d'une image dite "noire" (N) nous fournissent les niveaux planchers dans chaque canal (biais), celles tirées d'une image dite "blanche" (B) donnent les gains relatifs.

Pour un canal k donné, la mesure est corrigée de la manière suivante:

$$\tilde{\chi}_k = \frac{(\chi_k - N_k)}{(B_k - N_k)/max(B-N)}. \tag{4.1}$$

Cette valeur sera appellée, selon le canal, la réponse en cosinus m, en sinus m, ou l'ordre zéro ($\tilde{\chi}_k \in [C_{nm}, S_{nm}, C_{n0}]$). Les corrections sont simples et rapides à réaliser numériquement.

4.2.3 Calcul des signatures CHC

A la section 2.3, une technique de caractérisation à l'aide du système optique a été présentée. Une fois cette opération menée, les paramètres $\alpha, A_c, A_s, B_c, B_s$ sont connus. Le calcul des CHC est donc obtenu à l'aide des relations déduitent à la section 2.3.1:

$$\begin{aligned} H_{nm}(\alpha, A_c, A_s, B_c, B_s) = & \left(\frac{2}{C_{n0} A_c A_s \cos(m\alpha)}\right)^2 \Big[A_s^2\left(U_{nm}^c - B_c\right)^2 \\ & -2A_c A_s\left(U_{nm}^c - B_c\right)\left(U_{nm}^s - B_s\right)\sin(m\alpha) + A_c^2\left(U_{nm}^s - B_s\right)\Big] \end{aligned} \tag{4.1}$$

où

$$\begin{aligned} U_{nm}^c(f;\phi - \alpha/2, A_c, B_c) &= C_{nm}(f;\phi - \alpha/2, A_c, B_c) - \frac{1}{2} C_{n0}(f,\phi) \\ U_{nm}^s(f;\phi + \alpha/2, A_s, B_s) &= S_{nm}(f;\phi + \alpha/2, A_s, B_s) - \frac{1}{2} C_{n0}(f,\phi) \end{aligned} \tag{4.1}$$

Lors du calcul, les disparités résiduelles de niveau et de gain sont tenues en compte, ainsi que les défauts d'alignement angulaires relatifs des rosettes en cosinus et en sinus de chaque ordre.

4.2.4 Classification par réseaux de neurones

L'étape de classification par réseaux de neurones vise à fournir au système une robustesse face aux fluctuations statistiques de la signature CHC d'une cible, ainsi que la capacité de reconnaître cette cible sous différentes attitudes (rotations hors plan).

En principe, un réseau de neurones de structure hiérarchique convenablement entrainé, c'est-à-dire deux ou plusieurs perceptrons à une couche cachée, séparera les signatures provenant des différentes cibles en autant de classes.

On trouvera au chapitre suivant, les résultats de simulation numérique qui illustrent les performances atteignables à l'aide de cette méthode. Le lecteur désirant avoir plus d'information sur les réseaux de neurones et les structures hiérarchiques des réseaux de neurones est invité à consulter le rapport de la première partie du projet.

CHAPITRE 5

Résultats expérimentaux

Ce chapitre fait état des différents résultats expérimentaux obtenus à l'aide des méthodes d'extraction des CHC pour la reconnaissance des objets. La première section présente les résultats obtenus lors des expériences numériques réalisées au cours de la seconde partie du projet. Ces résultats indiquent que le comportement des CHC est conforme à celui prévu théoriquement en permettant l'invariance sous rotation dans le plan et sous échelle. Les tests effectués ont aussi permis de démontrer l'utilisation des réseaux de neurones pour la classification de modèles réduits de véhicules dans un contexte de reconnaissance 2D et dans un contexte particulier de la reconnaissance 3D.

Dans une seconde partie, les résultats optiques les plus récents sont présentés. Ces résultats ont tous été obtenus à l'aide du montage comportant deux ACL, un premier pour afficher les images et un second pour afficher les rosettes. Ces résultats confirment bien que le comportement global des CHC correspond à celui des calculs numériques. Selon les expériences réalisées, les CHC angulaires (CHCA) démontrent une meilleure capacité de discrimination que les CHC. Bien que d'autres travaux seraient nécessaires, les résultats observés sont encourageant quant à l'utilisation des CHC et CHCA comme outils pour la reconnaissance invariante d'objets.

5.1 Résultats numériques

Lors de la première partie du projet, des tests ont été effectués sur un nombre restreint d'images. Le rapport entre le nombre d'images tests (10) et le nombre d'images d'entraînement (100) étaient relativement faible . Les tests n'avaient portés alors que sur deux véhicules. Dans cette seconde partie les CHCA ont été introduits comme nouvel outil de reconnaissance. Les images ont été modifiées et le nombre d'images tests a été augmenté. De

plus, les signatures ont été extraites directement à partir des objets plutôt qu'à partir de leur transformée de Fourier (TF). Une seconde série de tests numériques a donc été entreprise afin de vérifier les capacités des CHCA pour la reconnaissance et afin de confirmer celles des CHC sur une nouvelle série d'images d'entraînement et de test.

5.1.1 Classification 2D

La qualité de la classification dépend de deux critères: la reconnaissance et la discrimination. La reconnaissance mesure le degré d'invariance de la technique face au différentes variations d'un même objet. La discrimination mesure quant à elle la capacité de différentier deux objets d'identité distincte. L'invariance aux transformations géométriques bidimensionnelles est une propriété intrinsèque des signatures CHC. C'est-à-dire, que les propriétés intrinsèques des CHC permettent l'invariance sous rotation dans le plan, échelle et intensité. Les tests de reconnaissance bidimensionnelle permettent d'évaluer le comportement original des CHC, à savoir si les mesures correspondent bien à la théorie.

5.1.1.1 Signatures CHC et CHC angulaires

Les expériences de reconnaissance ont été réalisées sur différentes images de modèles réduits de véhicules militaires. Les modèles utilisés sont un Jeep, un APC, un FAV et un Char. Les images originales étaient de dimensions élevées (640x480) par contre les tests en reconnaissance ont été effectués sur des images de dimensions réduites où la taille de l'objet est d'environ 180 pixels au maximum du grossissement (grossissement nominal unitaire x 2.0). Les signatures CHC ont été évaluées pour une rotation dans le plan de l'image du véhicule sur une plage de 0 à 360 degrés, par saut de 10 degrés, et pour des facteurs d'échelles variant de 0.5 à 2.0. Les rotations et les changements d'échelle ont été obtenus par échantillonnage sur les images originales. Ces facteurs correspondent à des objets de dimensions variant entre 45 pixels et 180 pixels pour une variation totale du facteur d'échelle de 4 fois. Les images, prises selon des vues de côté et de haut ont été traitées comme correspondant à des objets différents. Les images originales sont présentées à la figure 5.1.

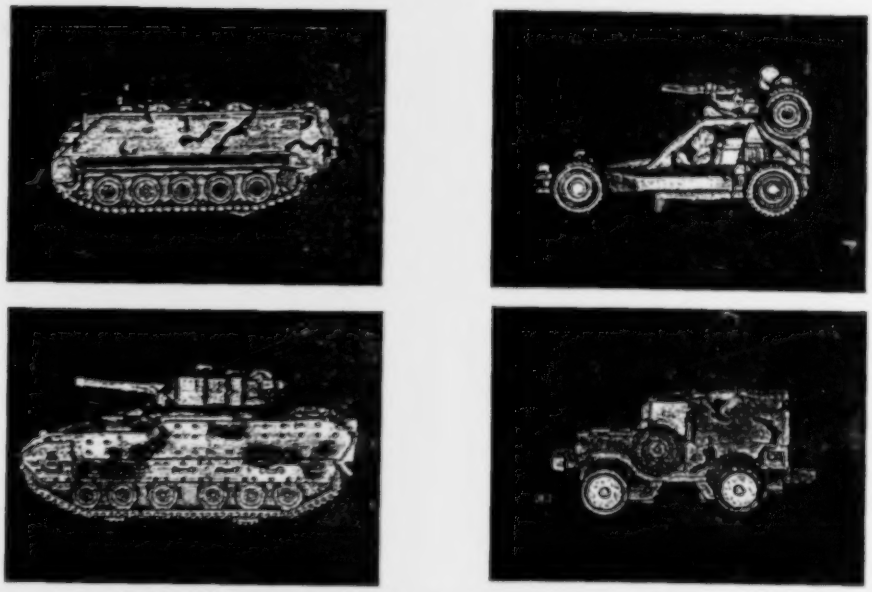


FIGURE 5.1 : a) Images des véhicules, vues de côté.

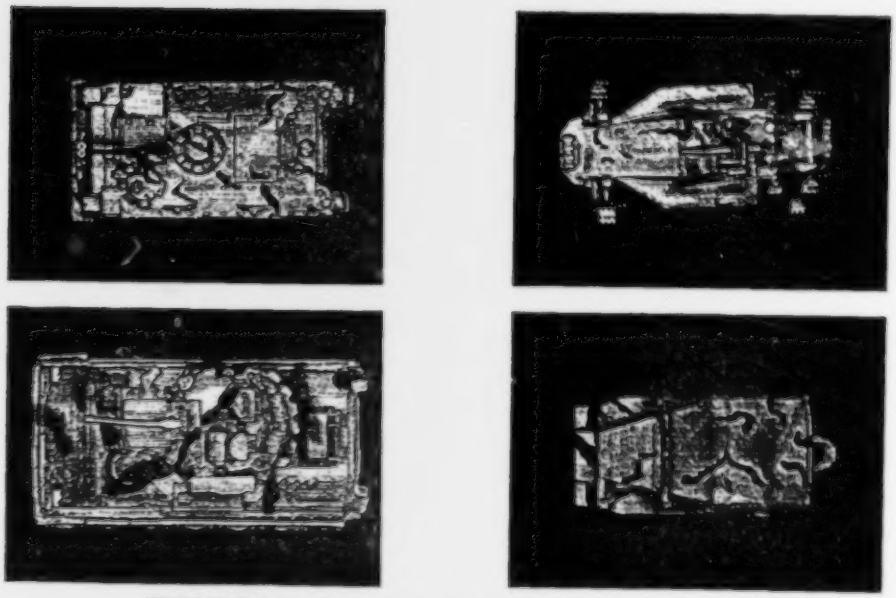


FIGURE 5.1 : b) Images des véhicules, vues de haut.

La figure 5.1 présente tout d'abord les signatures des 8 images tests pour un facteur d'échelle unitaire (1.0) et un angle de rotation nulle (0 degré). Les signatures CHC sont présentées sous forme d'histogramme pour chacun des objets. Chaque barre verticale correspond à la valeur du CHC pour l'ordre harmonique correspondant. On remarque que les signatures comportent, en général, une forte composante du second ordre harmonique. Ceci est explicable morphologiquement par une symétrie de base d'ordre 2 pour les véhicules, l'image étant centrée par rapport au centre géométrique de l'objet. Selon ces courbes il apparaît que l'aspect des signatures varie selon l'objet analysé. C'est cette variabilité interclasse des signatures qui permettra au réseau de neurones de distinguer les différents objets. Ainsi par exemple le FAV, qui est de forme plutôt triangulaire se distingue des autres par sa forte composante d'ordre 1.

Tel que prévu par la théorie, la valeur des coefficients décroît avec l'ordre des harmoniques. L'information contenue dans l'ordre zéro correspond à l'énergie totale et sert à la normalisation. L'ordre zéro n'est donc pas utilisé directement dans les tests de reconnaissance et n'apparaît pas sur les signatures présentées.

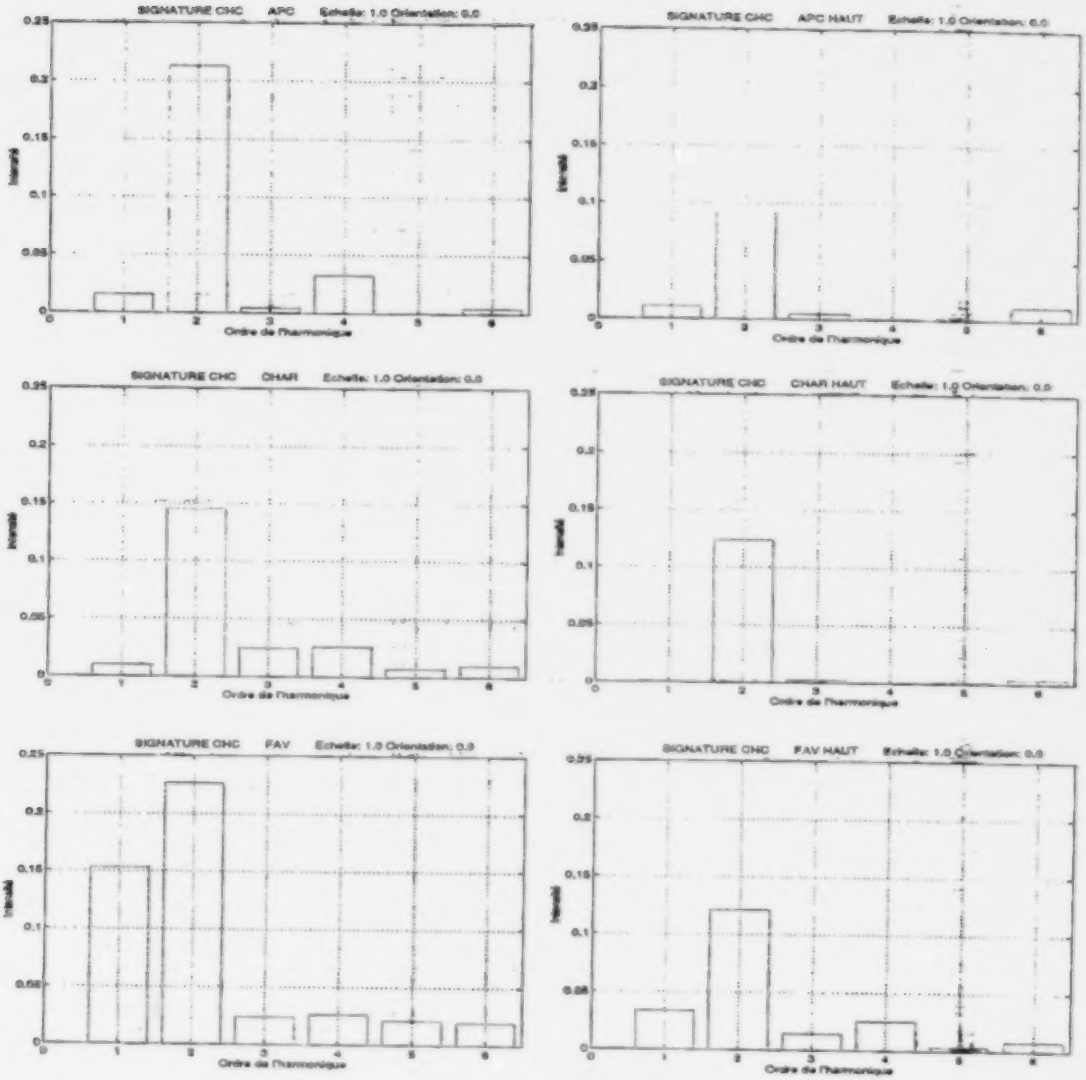


FIGURE 5.2 : Valeurs des CHC pour les images des APC, CHAR et FAV de la figure 5.1. Les graphiques de gauche correspondent aux vues de côté tandis que ceux de droite correspondent aux vues de haut.

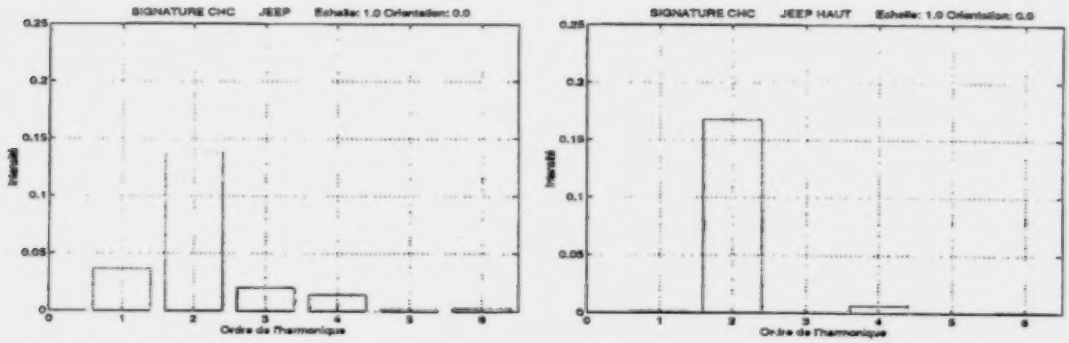


FIGURE 5.2 : Valeurs des CHC pour les images des JEEP de la figure 5.1. Les graphiques de gauche correspondent aux vues de côté tandis que ceux de droite correspondent aux vues de haut.

Les signatures CHC angulaires (CHCA) des images tests sont présentées aux figures 5.1 et 5.1. Puisque la signature fait référence aux paramètres de phase, la plage de variation des signatures s'étend entre 0 et 360 degrés. La valeur de la première harmonique, indicatrice de la rotation de l'objet, sert de référence angulaire pour les autres coefficients de la signatures. La première harmonique est donc fixée arbitrairement à 0 degré et la signature utilisée pour la reconnaissance comporte 5 coefficients. L'ordre zéro sert à la normalisation dans le calcul des signatures angulaires. On remarque que la variabilité interclasse des signatures CHCA est plus grande pour les différents véhicules modèles que la signature CHC. Ceci aura un impact positif sur le temps d'apprentissage du réseau de neurones.

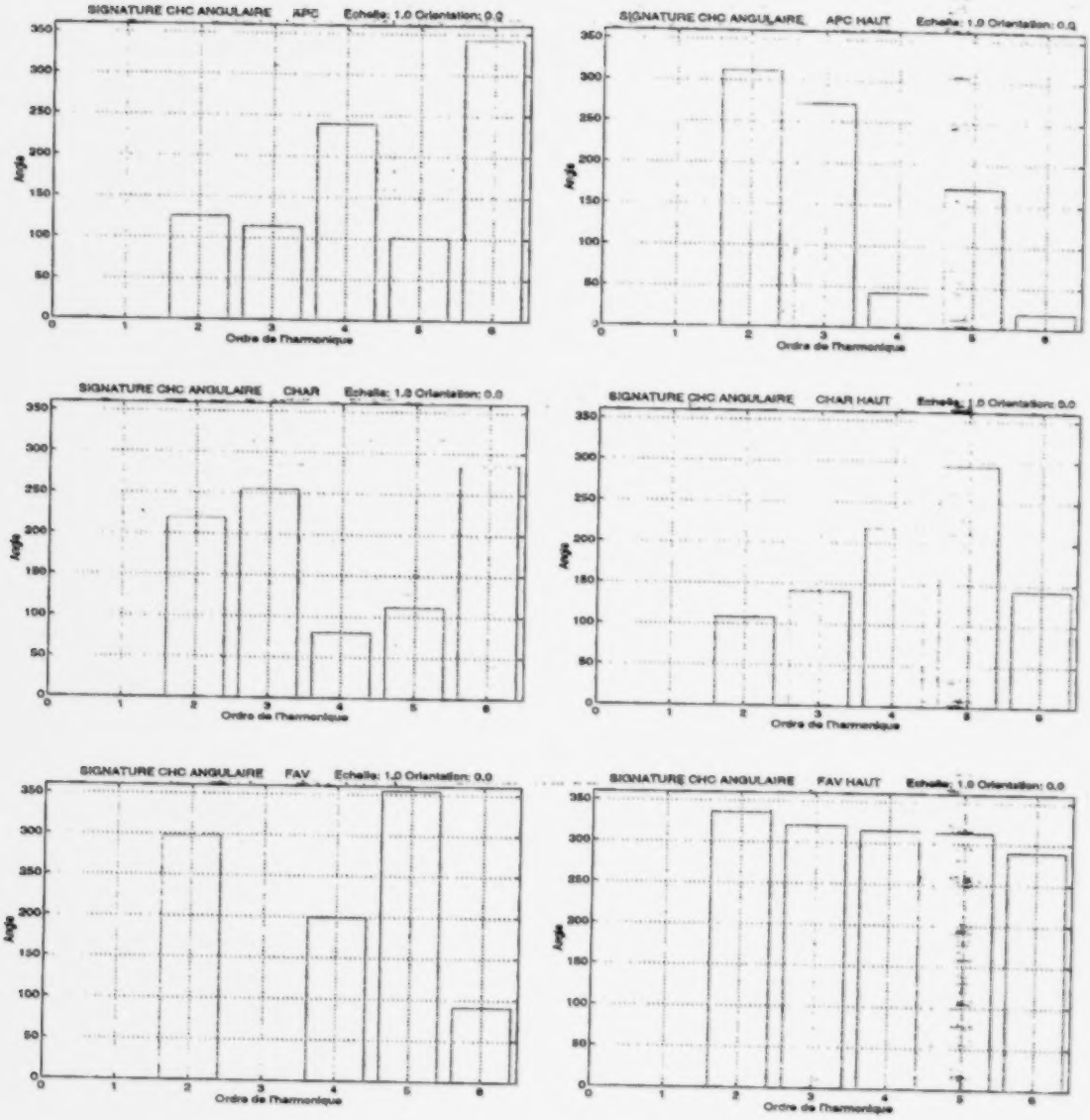


FIGURE 5.3 : Valeurs des CHCA pour les images des APC, CHAR et FAV de la figure 5.1.

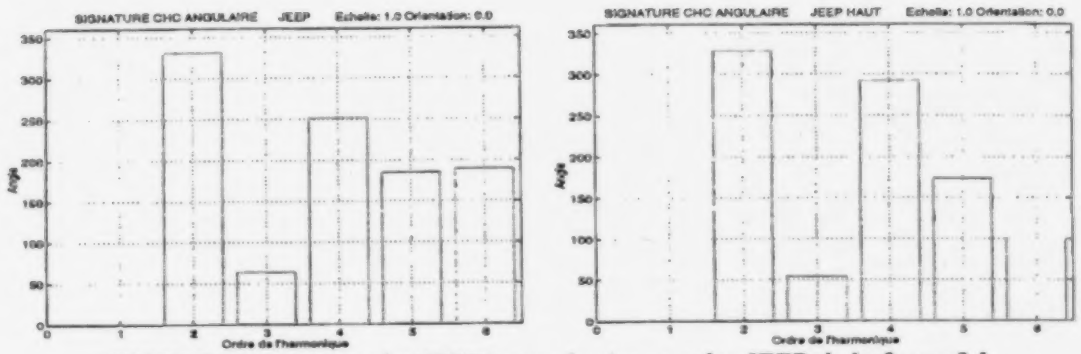


FIGURE 5.4 : Valeurs des CHCA pour les images des JEEP de la figure 5.1.

Les signatures CHC et CHCA sont invariantes sous rotation dans le plan et sous échelle. Ceci implique que les histogrammes présentés plus haut devraient rester inchangés quelque soit l'échelle ou la rotation dans le plan. Il est bon de vérifier l'invariance des signatures suite à une rotation ou à un changement d'échelle de la cible. La figure 5.1 présente les résultats d'une telle évaluation pour les coefficients 1 à 6 avec calcul de CHC pour le modèle du Char d'assaut. Chaque graphique représente la variation de la signature CHC pour une rotation dans le plan de l'objet sur une plage de 0 à 360 degrés. Pour chacun des graphes, chaque ligne correspond plus particulièrement à la variation de la valeur d'un coefficient pour la plage totale de rotation. Chacun comporte donc 6 lignes, correspondant aux 6 coefficients.

On remarque dans l'ensemble la stabilité des signatures. A la figure 5.1, on retrouve les graphiques pour les valeurs d'échelle de 0.5 à 2.0 par saut de 0.25. On remarque que les incréments les plus visibles en rotation sont dans le cas du plus faible facteur d'échelle, soit 0.5. De même, la variation du niveau du second coefficient est perceptible pour un changement du facteur d'échelle de 0.5 à 0.75. Cependant même à ce niveau de grossissement, la signature dans son ensemble est stable. De plus, il ne se produit pas d'interversion des niveaux des CHC des différents ordres. L'aspect de la signature CHC est donc conservé.

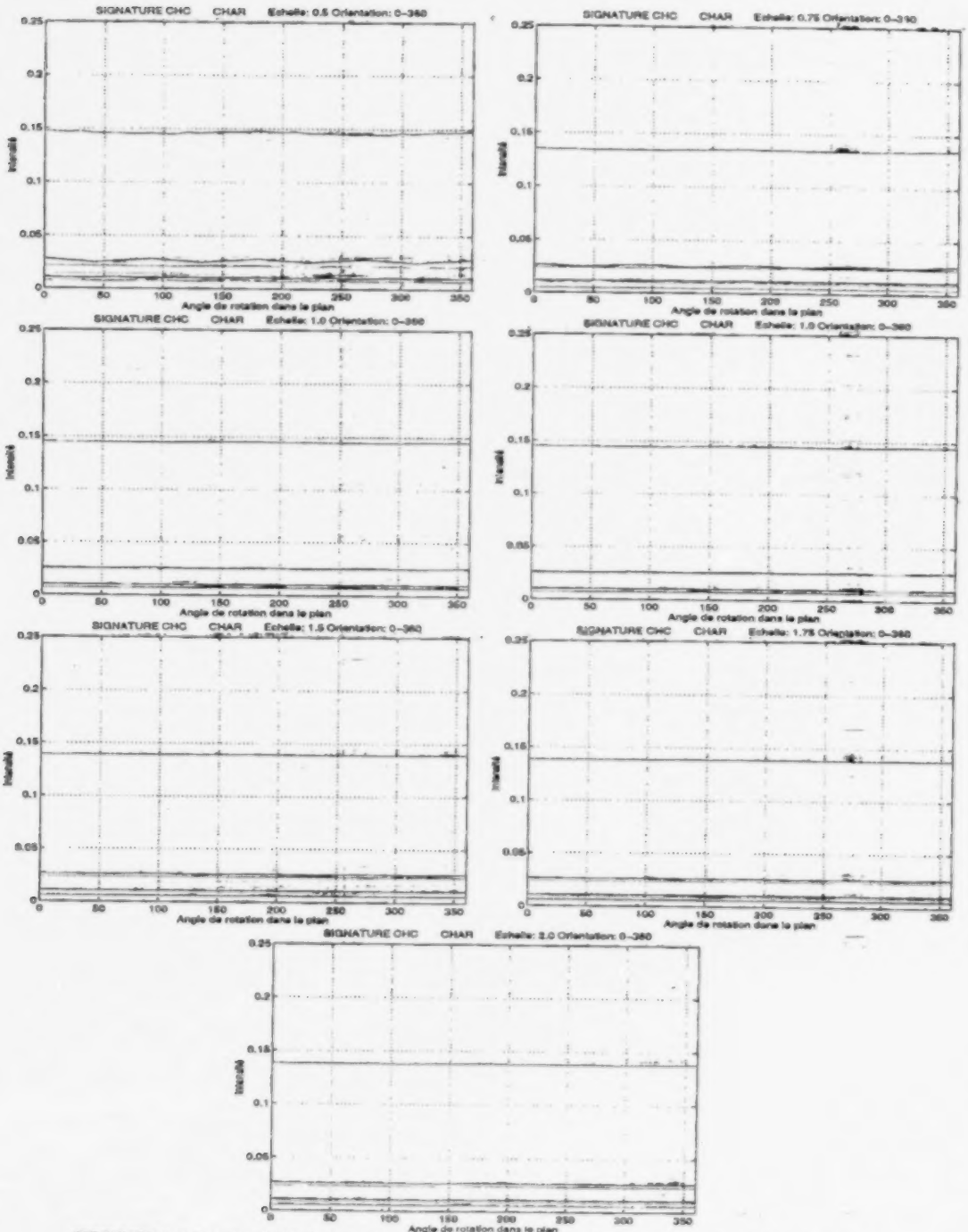


FIGURE 5.5 : Stabilité des signatures CHC pour la rotation dans le plan et le changement d'échelle des véhicules.

La figure 5.1 présente quand à elle l'invariance des signatures CHCA pour le modèle réduit du Char. Ici encore, la signature est stable et ne varie que très peu (~2-3%) sur les plages totales de rotation et d'échelle. On remarque pour le facteur d'échelle 0.5 un saut brusque de 360° pour l'une des composantes harmoniques. Ce phénomène est attribuable à la périodicité de l'angle. Cet artefact vient simplement de la référence de calcul et ne nuit pas aux tests de reconnaissances puisque l'on peut faire une transformation des données à l'aide de fonctions circulaires à l'entrée du réseau de neurones. On remarque que, comme pour les CHC, la signature CHCA, malgré certaines variations, conserve son aspect macroscopique pour les différents angles et échelles. Ce comportement est conforme à celui prévu par la théorie. On peut donc qualifier de faible variation globale dans le cas des invariances 2D. Les informations contenues dans les CHC et CHCA sont mutuellement indépendantes, elles peuvent donc être utilisées de concert ou individuellement pour une opération de reconnaissance.

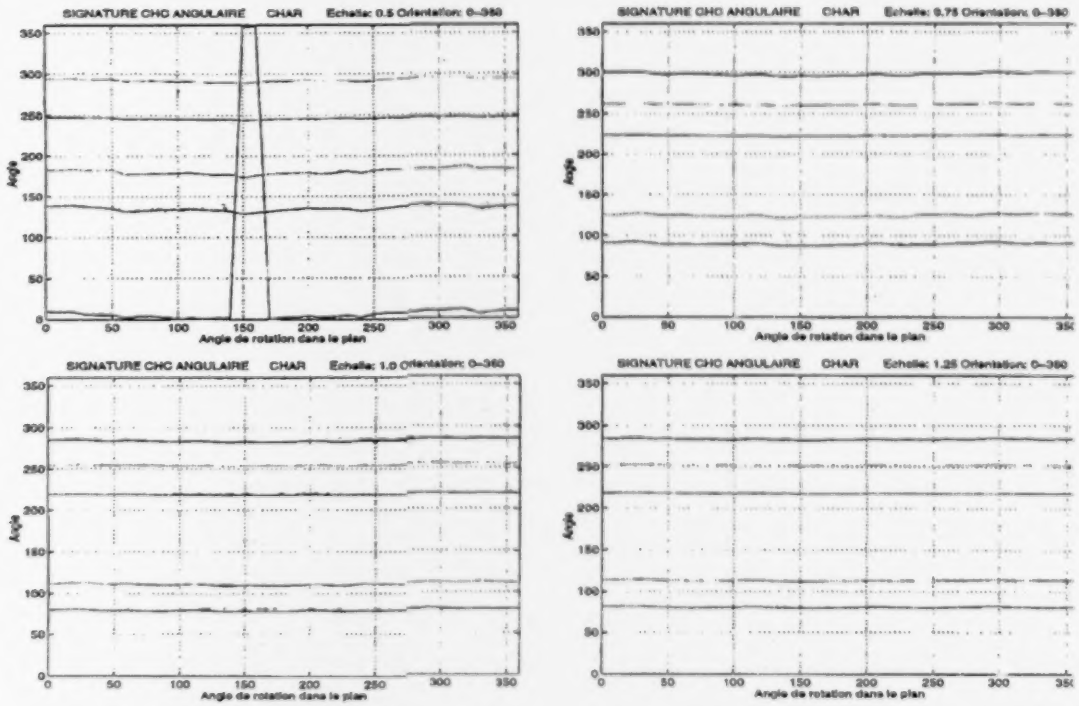


FIGURE 5.6 : Stabilité des signatures CHCA pour la rotation dans le plan et le changement d'échelle. Le saut d'angle au facteur d'échelle 0.5 correspond à une variation de 0 à 360 degrés et ne constitue pas une erreur.

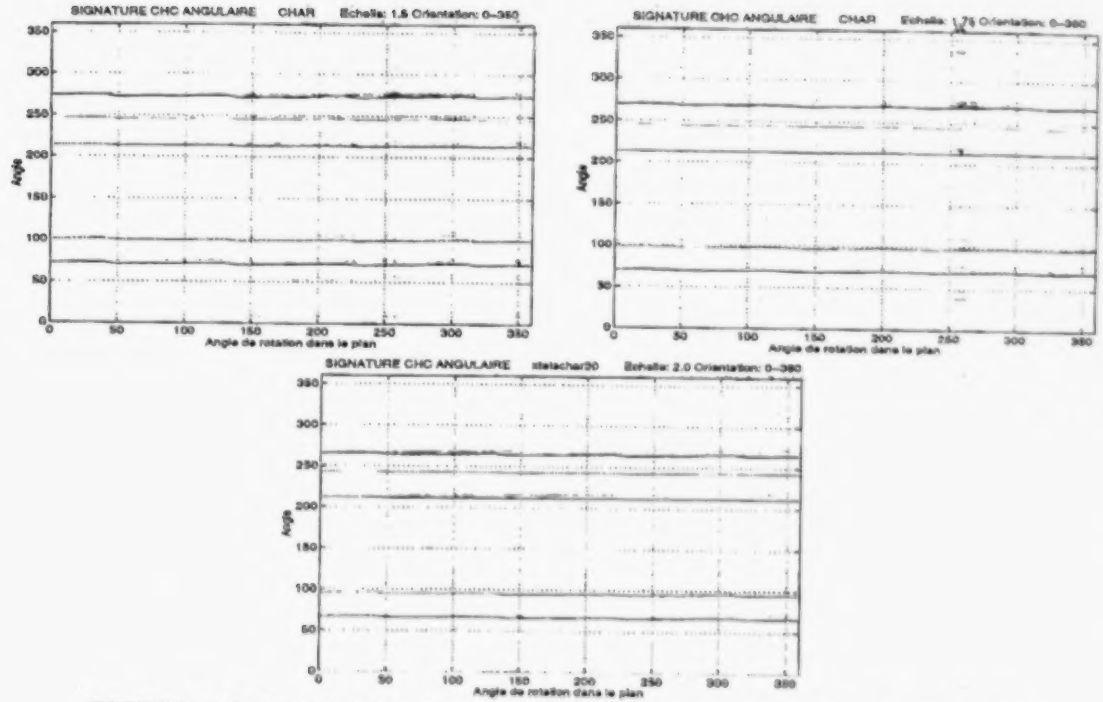


FIGURE 5.6 : Stabilité des signatures CHCA pour la rotation dans le plan et le changement d'échelle.

5.1.1.2 Reconnaissance et discrimination

Pour vérifier les capacités des CHC en reconnaissance et en discrimination, deux séries de tests ont été effectuées. Ainsi, les CHC et CHCA ont tous les deux été testés séparément. Les images de tests ont été créées à partir des rotations et changement d'échelle des huit images de modèles réduits présentées précédemment. Pour chaque image test, les signatures CHC et CHCA ont été calculées. Un nombre restreint de signatures ont été retenues pour l'apprentissage du réseau de neurones. Le réseau de neurones utilisé est un perceptron à une seule couche cachée. Toutes les signatures ont été présentées au réseau pour l'opération de test de reconnaissance. Le tableau 1 présente les caractéristiques et performances obtenues

TABLEAU 1. Caractéristiques et performances des réseaux neuronaux: 2D

CHC et CHCA NUMÉRIQUE 2D		
Paramètres du réseau	CHC	CHCA
# Neurones entrées	6	5
# Couches cachées	1	1
# Neurones Couches cachées	36	36
# Neurones sorties	8	8
# Images entraînement	18	8
# Images tests	2072	2072
# Itérations	7214	261
Reconnaissance	Complète	Complète
Discrimination	Complète	Complète

Tel que l'indique le tableau 3, la reconnaissance dans les deux cas est complète, c'est-à-dire que, pour les images testées, les deux réseaux de neurones ont été capables de reconnaître correctement les 2072 images tests. Dans le cas du CHC standard, 18 images ont dû être présentées lors de l'apprentissage pour obtenir une reconnaissance complète, c'est-à-dire environ 2 images par objet. Les images supplémentaires sont requises pour tenir compte des variations supplémentaires apparaissant à quelques facteurs de rotation ou d'échelle particuliers. Dans tous les cas, lorsque le réseau a reconnu un véhicule modèle, il a aussi pu discriminer les autres véhicules modèles. Les sorties, sur 8 neurones, ont été encodées de façon orthogonale. Ainsi, par exemple, une sortie "10000000" correspond au APC et une sortie "01000000" correspond au FAV, etc.

Lors de l'entraînement, dans le cas des CHC angulaires, un seul exemplaire de chaque objet à été nécessaire pour permettre l'identification de tous les véhicules modèles. Ici encore la discrimination a été complète. La principale différence entre les deux réseaux se situe au niveau du nombre d'itérations requises pour compléter l'apprentissage. Les CHCA ont requis 261 itérations contre 7214 pour les CHC. Ceci semble indiquer que la variabilité interclasse des signatures CHCA est plus élevée que celle des CHC, facilitant d'autant la tâche du réseau de neurones pour la discrimination des signatures. Il est intéressant de noter que les deux

réseaux ont permis, indépendamment, la reconnaissance et la discrimination de tous les objets. Il existe donc toujours une possibilité d'intégration des deux types d'informations.

5.1.2 Classification 3D

Au niveau de l'invariance tridimensionnelle, les variations associées à la rotation hors-plan sont tenues en compte à l'aide des réseaux neuronaux, et la performance dépend surtout des opérations d'apprentissage du réseau de neurones. À partir des signatures d'un véhicule modèle à différentes rotations hors-plan, on entraîne le réseau à reconnaître le véhicule modèle pour une grande diversité de points de vue. Chaque signature étant elle-même invariante dans le plan, la présence du réseau de neurones permet l'ajout d'une part d'invariance supplémentaire. S'il est vrai que les tests bidimensionnels évaluent le bien fondé de l'usage des CHC comme outil d'identification, les tests tridimensionnels évaluent surtout la capacité des réseaux neuronaux comme outil d'apprentissage et de reconnaissance. Ceci est dû au fait que l'information provenant de différents points de vue (projection 2D) de l'objet 3D est incomplète.

5.1.2.1 Signatures CHC et CHC angulaires.

Contrairement à ce qui prévaut dans le cas 2D, les signatures CHC ne sont pas invariantes pour une rotation hors plan d'un objet. Pour reconnaître complètement un objet, on doit pouvoir reconnaître les signatures correspondant aux différents points de vue de l'objet. Chaque vue étant compressée en un nombre restreint de coefficients CHC, il en résulte des réseaux de neurones de taille réduite même si l'on doit tenir compte de plusieurs vues différentes pour chaque véhicule. Pour réaliser les expérimentations de reconnaissance 3D, les quatre véhicules modèles précédents ont été observés en vue latérale sur 360 degrés de plage de rotation par incrément de 15 degrés. Au total 24 vues pour chaque véhicule modèle ont été utilisées. L'apprentissage est quelque peu différent pour le cas 2D. La signature CHC est plus stable sous rotation hors axe mais montre une variabilité interclasse moins prononcée que les CHCA. De ce fait la signature CHC varie lentement avec la rotation hors-plan. De façon générale, la signature CHC demande un ensemble d'entraînement plus restreint. Les vues doivent toutefois être classées hiérarchiquement avant l'identification finale ce qui requiert plusieurs réseaux de neurones. Les signatures CHCA montrent quant à elles une plus grande plage de variation interclasse. L'ensemble d'entraînement est plus grand mais un seul

réseau est requis. De plus, le classificateur neuronal pour les signatures CHCA requiert un nombre restreint d'itérations.

5.1.2.2 Discrimination et reconnaissance

Pour la reconnaissance par signature CHC, les données ont été classées en deux étapes. Dans un premier temps, un premier réseau identifie l'appartenance du véhicule modèle présenté à la classe Face ou Côté. L'information est par la suite transférée au réseau de neurones correspondant pour identifier le type de véhicule modèle présent à l'entrée du système.

TABLEAU 2. Caractéristiques et performances des réseaux de neurones: 3D CHC

CHC NUMÉRIQUE 3D			
Paramètres du réseau	Valeurs Face-Côté	Valeurs Face- Identification	Valeurs Côté- Identification
# Neurones entrées	6	6	6
# Couches cachées	1	1	1
# Neurones couches cachées	10	10	12
# Neurones sorties	2	4	4
# Images entraînement	21	11	14
# Images tests	96	16	80
# Itérations	1112	28528	16328
Reconnaissance	Complète	Complète	97%
Discrimination	----------	Complète	97%

Le tableau 2 présente les paramètres des trois réseaux utilisés pour la classification des véhicules modèles à l'aide des signatures CHC. On remarque tout d'abord, en combinant les résultats, que les réseaux ont pu effectuer un taux de reconnaissance de plus de 97% sur les images considérées. Le nombre d'itérations nécessaires à la convergence du réseau varie entre 1000 et 30000 itérations. Le nombre d'itérations est un reflet de la plage de variation restreinte des CHC. Les réseaux sont toutefois de dimensions forts réduites puisque le nombre maximal

de neurones dans la seule couche cachée requise est de 12 neurones. Ce test reflète l'intérêt de la méthode pour la reconnaissance 3D. Un nombre plus élevé d'images serait utiles pour une évaluation à grande échelle de la méthode. Il est à noter que la classification Côté-Face est flexible. Ainsi une image située à la frontière entre les deux classes pourrait se voir classer dans une ou l'autre des classes selon le véhicule modèle ou la vue. Cependant, lorsque ces données sont transférées au réseau suivant, elle peuvent être classées correctement étant donné l'apprentissage du réseau de neurones. Ainsi, dans les tests réalisés, sur les 16 images transférées au réseau Face, 13 ont été identifiées correctement par le réseau Côté.

Les résultats pour les signatures CHCA sont données au tableau 3. Il est à noter qu'un seul réseau a été utilisé pour la classification. Ceci est un reflet de la plus grande plage de variation des signatures d'un objet à l'autre. Les signatures, pour un véhicule modèle donné, affichent une plus grande variation en fonction d'une rotation de l'objet sur 360 degrés mais sont par contre plus discriminantes par rapport aux autres véhicules modèles.

TABLEAU 3. Caractéristiques et performances du réseau neuronal: reconnaissance 3D avec CHCA

CHCA NUMÉRIQUE 3D	
Paramètres du réseau	Valeurs
# Neurones entrées	10
# Couches cachées	1
# Neurones Couches cachées	28
# Neurones sorties	4
# Images entraînement	47
# Images tests	96
# Itérations	185
Reconnaissance	Complète
Discrimination	Complète

On remarque tout d'abord le taux élevé de reconnaissance puisque le réseau est à même de distinguer toutes les images présentées. Le réseau requiert davantage d'images d'entraînement, mais un nombre plus petit d'itérations (185). Ceci témoigne de l'orthogonalité et des capacités de discrimination intrinsèques des signatures CHCA. La taille

du réseau est aussi restreinte avec 28 neurones dans la seule couche cachée. En moyenne, une vue du véhicule modèle à tous les 30 degrés à été requise pour un apprentissage complet. Les signatures angulaires ont été transformées en fonctions circulaires sinus et cosinus, ce qui explique la présence de 10 neurones à l'entrée. Des valeurs de sortie orthogonales pour les quatre véhicules ont été attribuées aux quatre neurones de sortie. Ici encore, des tests sur un nombre plus élevées d'images permettraient une évaluation plus précise des performances de la technique, en particulier en présence de bruit.

5.2 Résultats optiques

L'extraction des CHC peut être réalisée numériquement ou optiquement. L'avantage de cette dernière solution étant le gain en vitesse et la possibilité d'analyser les images présentées au plan d'entrée en temps réel, c'est-à-dire à la cadence vidéo. Les résultats présentés ci-dessous ont été obtenus à l'aide du montage à deux ACL.

Les expérimentations optiques ont été réalisées à l'aide des quatre images de véhicules modèles en vue latérale. Les images originales ont été réduites afin que leur taille soit de l'ordre de 180 pixels. Les quatre images des véhicules modèles ont été tournées dans le plan sur une plage de 360 degrés par saut de 20 degrés. Les images tournées et à différentes échelles sont générées numériquement par échantillonnage des images originales. De plus, des images ayant des facteurs d'échelle de 0.4, 0.6, 0.8 et 1.0 ont été générées. Lors des expériences, les images ont été présentées tour à tour à l'entrée du système optique. Le système optique utilisé est celui à 2 ACL puisque le système à réseau de diffraction n'a pas donné les résultats escomptés. Pour chaque image, les 13 rosettes représentant les 6 premiers ordres des harmoniques de même que l'ordre zéro ont été affichés sur le second ACL. De plus, les niveaux d'énergie correspondant à une entrée uniforme blanche (transparente) et une entrée uniforme noire (opaque) ont été mesurés afin de calibrer les niveaux d'énergie. Les valeurs mesurées ont par la suite été analysées pour calculer les CHC et CHCA. Il est à noter que les deux signatures proviennent exactement de la même série de mesures sur lesquelles sont simplement effectuées des traitements différents.

5.2.1 Reconnaissance et discrimination 2D

Les tests en reconnaissance 2D ont été effectuées sur 96 images représentant des plages de rotation dans le plan de 0 à 360 degrés et des facteurs d'échelle de 0.4 à 1.0.

Une mesure de l'illumination est tout d'abord enregistrée. A l'aide de cette mesure une figure d'uniformisation est calculée et superposée à chaque rosette afin d'obtenir une illumination uniforme sur toute la surface des images affichées. De plus, les rosettes sont modifiées par rapport à leur courbe analytique pour tenir compte de la non-linéarité des ACL. Cette étape n'est effectuée qu'une seule fois lors de l'initialisation du système et n'a pas à être réalisée de nouveau en mode d'opération du système. La figure d'uniformisation de référence est identique pour toutes les harmoniques. La figure 5.1 présente les figures d'illumination avant uniformisation à gauche et après uniformisation à droite. L'éclairement est plus constant après l'uniformisation qu'avant cette dernière. Les franges apparaissant sur l'image sont simplement dues au moiré créé entre la structure de la caméra CCD et celle des ACL. Le moiré n'a pas d'incidence sur les mesures puisqu'un détecteur simple est utilisé dans le montage.

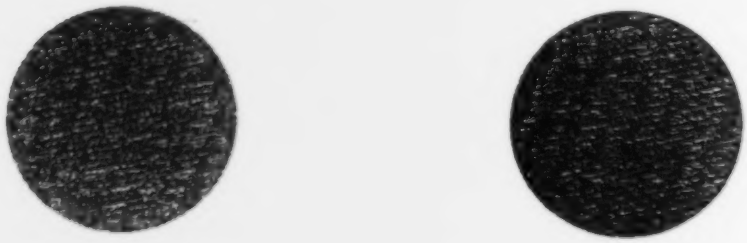


FIGURE 5.7 : Uniformisation de la figure d'illumination. Les franges de moiré sont crées par la caméra et l'ACL. Celles-ci ne sont pas présentes lors des mesures.

Les étapes sont simples pour reconnaître un véhicule modèle donné. L'image est multipliée optiquement par chacune des rosettes et un détecteur mesure l'énergie lumineuse contenue dans le produit des images. Chaque rosette génère donc une valeur scalaire. De ces

valeurs sont extraites les CHC et CHCA de façon numérique. Ce calcul nécessite peu d'opérations et est par conséquent rapide. Les coefficients CHC et CHCA sont par la suite transférés au réseau de neurones, lequel en phase d'apprentissage adapte ses poids aux objets à reconnaître (entraînement), alors qu'il procède à l'identification en phase de reconnaissance (opération et tests).

La figure 5.1 présente la première étape de la reconnaissance d'un objet. Lors de cette étape, les objets sont affichés sur le premier ACL alors que les rosettes sont affichées sur le second ACL. Les images présentées montrent tout d'abord le transporteur de troupe APC multiplié par une rosette cosinusoïdale d'ordre 6 et le même transporteur multiplié par une rosette sinusoïdale d'ordre 6. La seconde série d'images présente les rosettes d'ordre 5 multipliant un char d'assaut incliné de 60 degrés. Dans une opération de reconnaissance,

treize rosettes sont présentées à tour de rôle pour les ordres variant de zéro à six. Le montage optique calcule les valeurs des C'_{nm} et S'_{nm} (voir section 2.1).

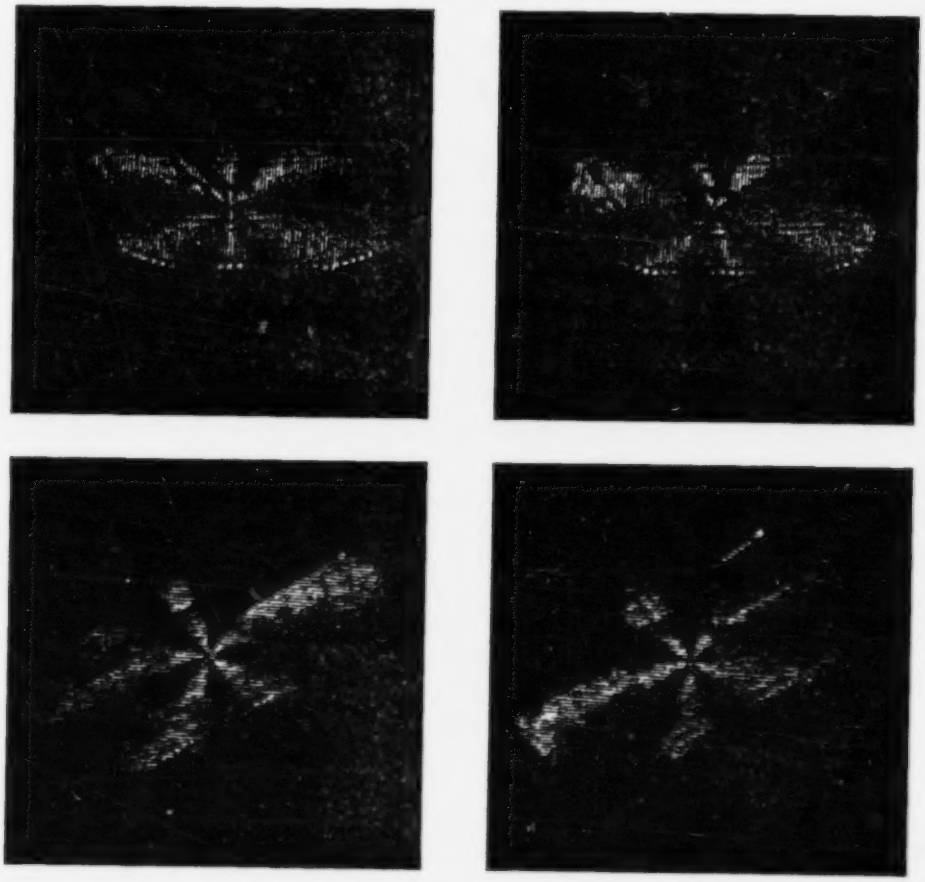


FIGURE 5.8 : Images résultant de la multiplication optique des véhicules par les rosettes sinusoïdales et cosinusoïdales.

5.2.1.1 CHC angulaires

Le signal obtenu sur le détecteur pour les rosettes cosinusoïdales et sinusoïdales avec à l'entrée le CHAR est présenté en figure 5.9. Les courbes présentées montrent l'évolution du signal mesuré sur le détecteur pour une rotation de l'objet sur une plage de 360 degrés par

saut de 20 degrés. Les points identifiés par des ronds et situés à l'extrême droite de la courbe correspondent au changement d'échelle de l'image pour des facteurs variants de 0.4 à 1.0.

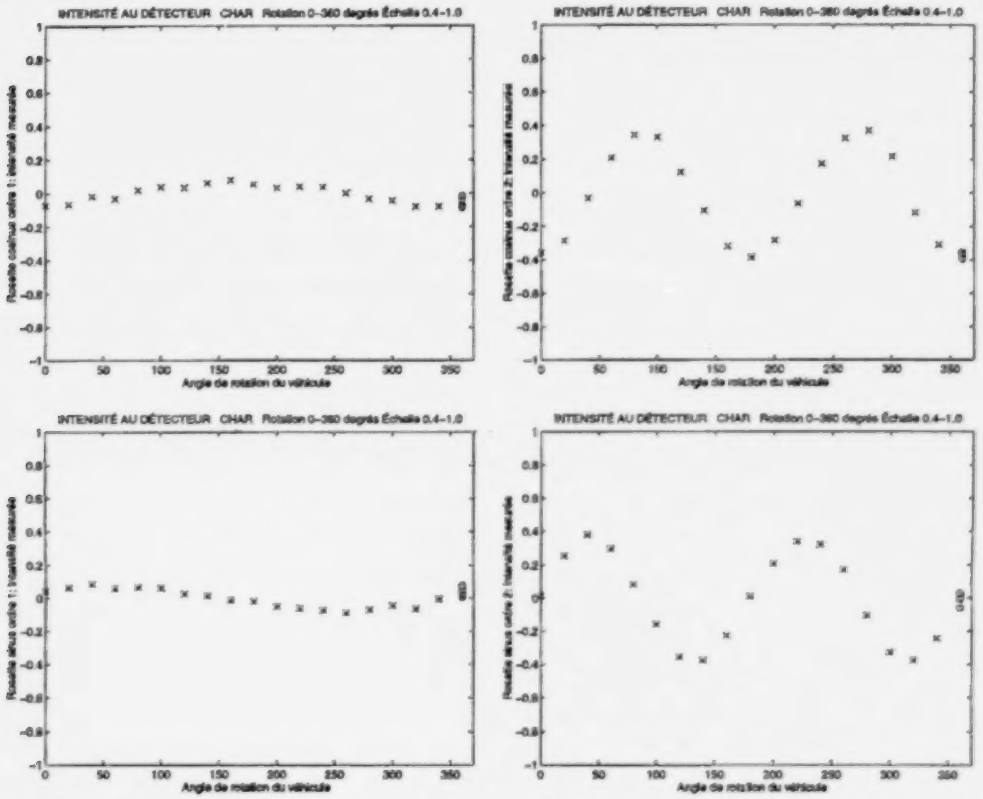


FIGURE 5.9 : Réponse obtenue sur le détecteur pour les deux premiers harmoniques couplées à un CHAR.

La figure 5.10 présente les mêmes résultats pour le modèle Jeep. Comme le montre ces figures, l'amplitude des courbes cosinusoïdales varient selon les objets. C'est toutefois dans le déphasage relatif entre les différents ordres que se retrouve l'information permettant la

génération des CHCA, cette information est en fait une indication de l'alignement relatif des composantes harmonique des différents ordres constituant l'objet.

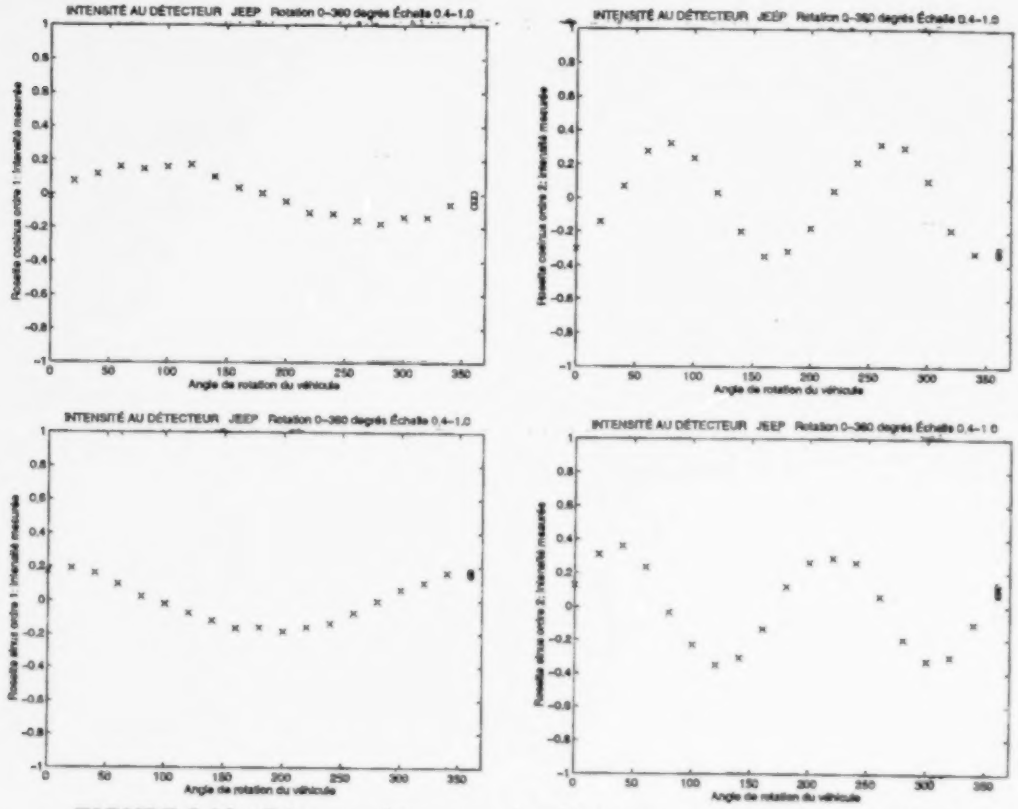


FIGURE 5.10 : Réponse obtenue sur les détecteurs pour les deux premières harmoniques couplées à un Jeep.

Comme le montre les figures précédentes, les déphasages relatifs entre les différents ordres CHC diffèrent d'un objet à l'autre. Cette information est caractéristique de l'objet et constitue l'essence de la signature CHCA. On remarque que les courbes ne comportent pas de discontinuités et sont dans l'ensemble stables. Les CHCA sont extraits en calculant pour chaque point la phase déterminée par les courbes sinus et cosinus. Toutefois, puisque la signature tourne avec l'objet il est nécessaire de référencer la signature. Pour ce faire, la seconde harmonique conjointement avec l'information de la première harmonique, est utilisée comme référence pour la signature CHCA. Ce faisant les six ordres mesurés donnent

naissance à cinq paramètres indépendants permettant la génération de la signature CHCA. L'information de l'ordre deux n'est toutefois pas perdue puisqu'elle permet d'obtenir une indication de la rotation dans le plan du véhicule présenté à l'entrée du système de reconnaissance. La figure 5.1 présente l'évolution de l'information angulaire sur la rotation des différents véhicules modèles évaluée à l'aide de l'information contenue dans la seconde et la première harmonique CHCA.

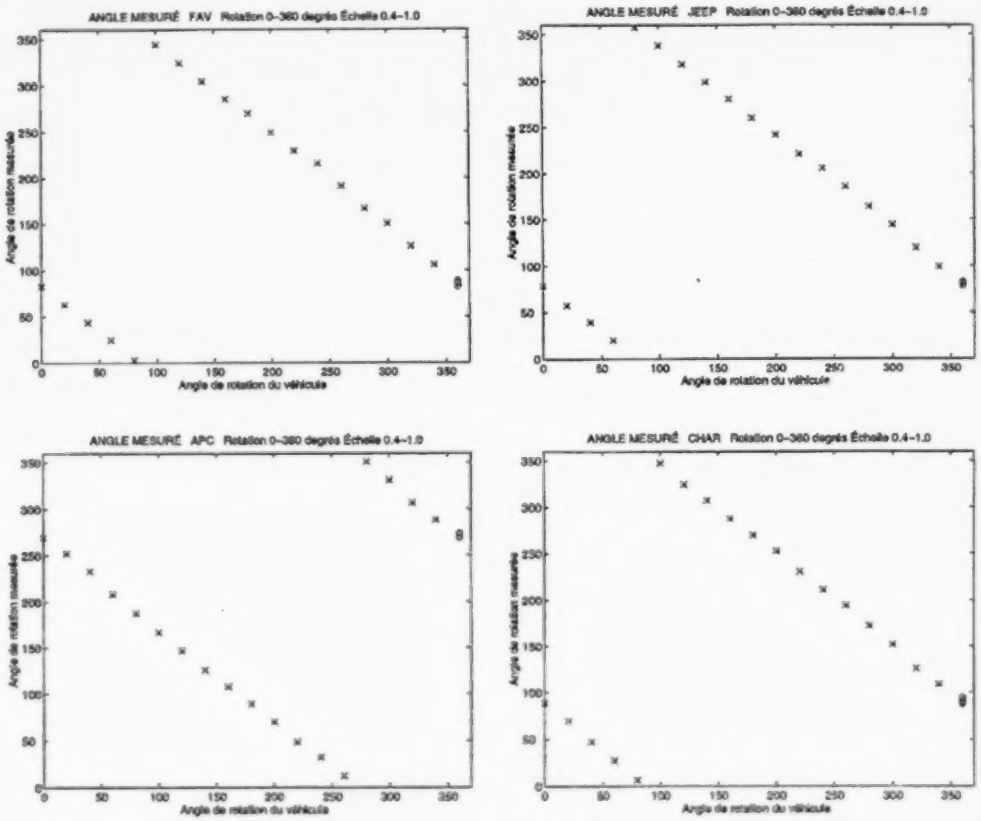


FIGURE 5.11 : Évaluation de la rotation des véhicules à partir des données optiques mesurées en comparaison avec angle réel de rotation du véhicule.

Chaque véhicule est représenté par un graphique pour des rotations variant de 0 à 360 degrés. L'origine dépend de chaque objet. On remarque que les courbes donnent une indication linéaire et stable de l'orientation de l'objet. Un certain gain dans l'évaluation de l'orientation est réalisé en couplant l'information de la seconde harmonique à celle de la

première harmonique. La seconde harmonique tourne deux fois plus rapidement que la première harmonique mais possède l'avantage de montrer une meilleure profondeur de modulation et une résolution plus élevée. Pour distinguer entre un angle du premier hémicercle de rotation ou du second hémicercle l'information de la première harmonique est utilisée. Ainsi, pour chaque variation de 0 à 360 degrés de la première harmonique, la seconde varie de 0 à 720 degrés. En couplant l'information des deux harmoniques, l'angle est identifié de façon unique.

Il est à noter que la mise en référence de la signature CHCA pour les fins de détection ne nécessite aucune information a priori ni d'hypothèse sur l'objet présenté à l'entrée. La signature CHCA est simplement calibrée à l'aide de la seconde harmonique mesurée. La stabilité des signatures CHCA évaluée pour les différents véhicules modèles et les différents ordres sont présentées à la figure 5.12.

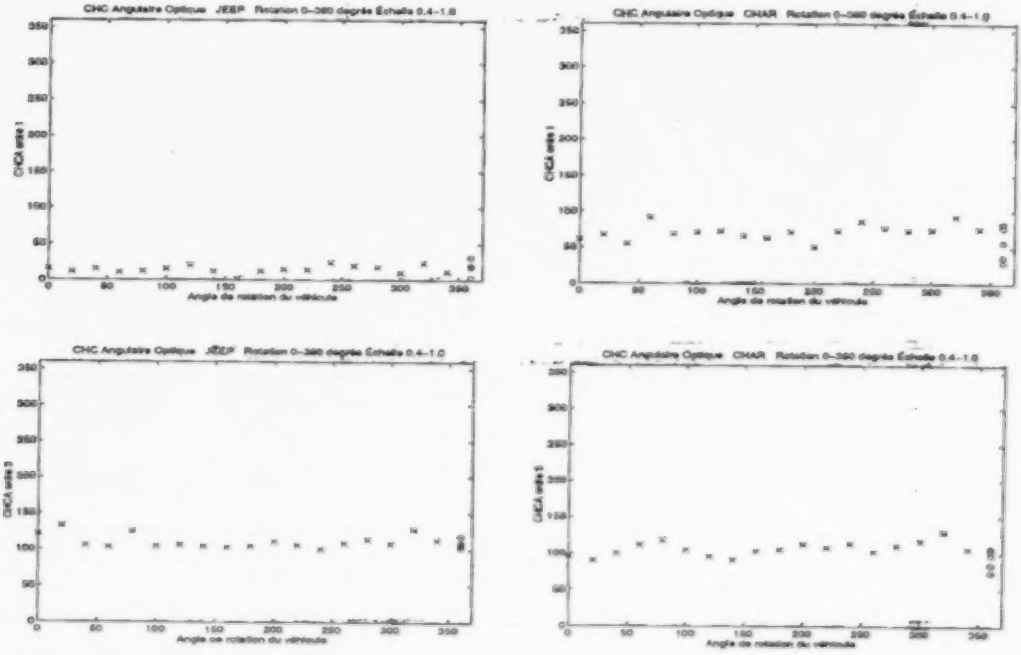


FIGURE 5.12 : Stabilité des CHCA optiques évaluées pour différents ordres pour le JEEP et le CHAR

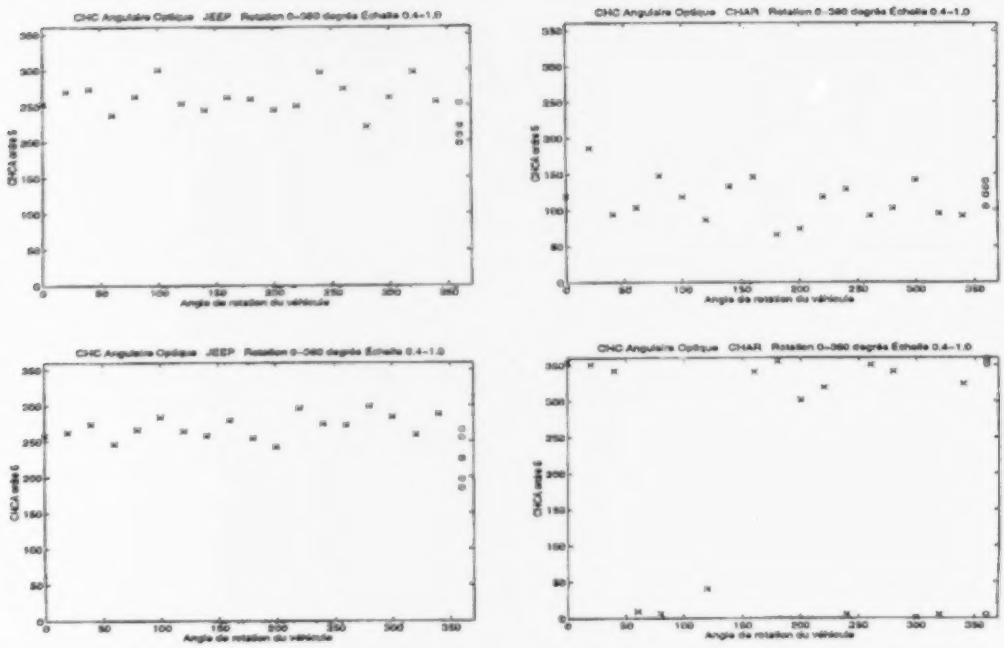


FIGURE 5.12 : Stabilité de CHCA optiques calculés pour différents ordres pour le Jeep et le Char.

On remarque tout d'abord que les signatures CHCA sont différentes pour les deux véhicules modèles présentés, les signatures présentent une certaines variabilité selon l'angle et l'échelle mais permettent néanmoins une séparation adéquate des objets. De façon générale, la signature est plus étendue pour les ordres supérieurs que pour les ordres inférieurs. Cependant, les ordres supérieurs sont bien bornés et même le sixième ordre montre une bonne stabilité considérant le peu d'énergie normalement contenue dans la valeur CHC. La capacité de discrimination est toutefois évidente et les signatures présentent une bonne orthogonalité telle que le prévoyait les simulations numériques sur les CHCA.

TABLEAU 4. Caractéristiques et performances du réseau neuronal: 2D CHCA optique

CHCA OPTIQUE 2D	
Paramètres du réseau	Valeurs
# Neurones entrées	10
# Couches cachées	1
# Neurones Couches cachées	28
# Neurones sorties	4
# Images entraînement	6
# Images tests	92
# Itérations	62
Reconnaissance	Complète
Discrimination	Complète

Les tests de reconnaissance ont été réalisés à l'aide d'un réseau de neurones de type perceptron à une seule couche cachée. Pour l'entraînement, 6 signatures optiques ont été présentées au réseau de neurones, lequel a convergé en 62 itérations. Les signatures optiques des 92 images analysées ont par la suite été présentées au réseau neuronal, lequel a reconnu convenablement toutes les images tout en affichant une discrimination complète pour toutes les reconnaissances effectuées. Ces résultats très encourageant sont regroupés au Tableau 4. Ils tendent à démontrer le potentiel des CHCA comme outil de reconnaissance. Des tests sur un plus grand nombres d'images permettraient une évaluation plus poussée des performances des CHCA.

5.2.1.2 CHC

Le calcul des CHC est réalisé à partir des rosettes des 6 premiers ordres. Les résultats de l'évaluation des CHC pour les trois premiers ordres de chacun des véhicules modèles sont présentés à la figure 5.13. Comme dans le cas des CHCA, les 'x' correspondent aux valeurs des CHC pour les rotations de l'objet de 0 à 360 degrés, alors que les 'o' correspondent aux valeurs CHC pour des changements d'échelle du véhicule modèle. Ici encore, les quatre véhicules modèles sont le JEEP, le CHAR, le APC et le FAV. On remarque tout d'abord que

les signatures sont dans l'ensemble uniformes, mais comportent une certaine variabilité. La variabilité interclasse des signatures pour différents véhicules modèles est, dans l'ensemble, moins grande que celle des CHCA. C'est-à-dire que les valeurs CHCA des différents ordres sont réparties plus uniformément sur la plage de variation admissible.

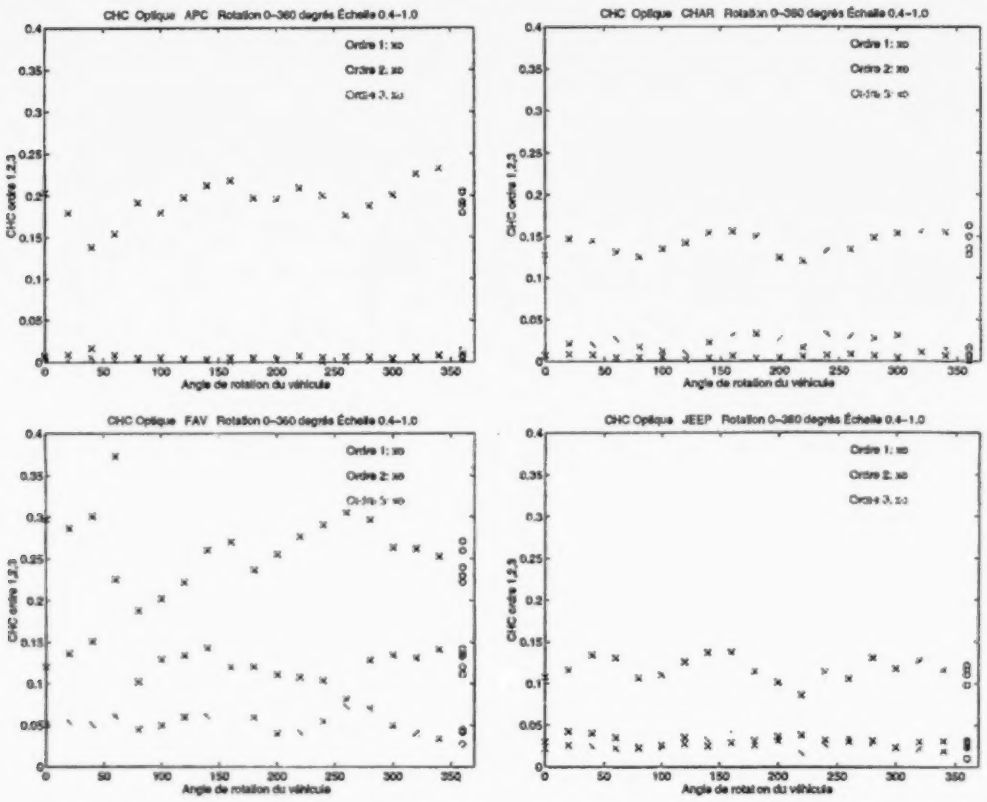


FIGURE 5.13 : Calcul de CHC optiques pour les ordres 1,2,3 pour 4 véhicules modèles différents: l'APC, le CHAR, le FAV et le JEEP.

Néanmoins, il existe des différences non négligeables entre les différents véhicules modèles. C'est le cas du FAV, lequel montre une composante d'ordre 1 ayant une valeur beaucoup plus élevée que les autres harmoniques. Cette valeur s'explique par la forme plus triangulaire du FAV. La figure 5.1 montre les résultats obtenus lors de l'extraction des CHC d'ordres 4, 5 et 6. On remarque tout d'abord qu'il y a dispersion des valeurs selon les angles

de rotation bien que certaines tendances dans les valeurs sont perceptibles. Leur influence sur la reconnaissance d'un objet sera donc limitée puisque les signatures sont moins discriminantes entre les différents véhicules modèles que celles des ordres inférieurs.

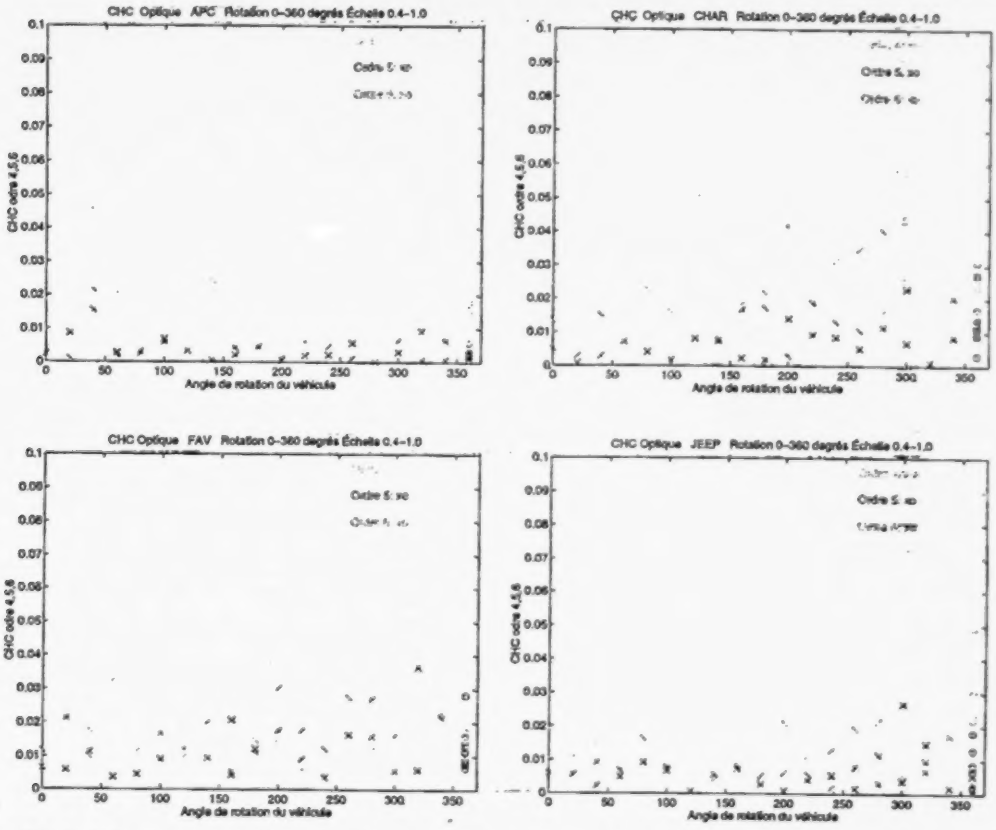


FIGURE 5.14 : Stabilité des signatures CHC optiques des ordres supérieurs pour quatre véhicules modèles: l'APC, le CHAR, le FAV et le JEEP

Les CHC recueillis ont été utilisés pour entraîner un réseau de neurones de type perceptron à une seule couche cachée. Au total, onze signatures ont été utilisées pour entraîner le réseau de neurones. Les tests ont été effectuées sur 92 images, dont 81 ont été utilisées uniquement lors des tests. Le tableau 5 présente les résultats obtenus.

TABLEAU 5. Caractéristiques et performances du réseau neuronal: reconnaissance 2D optique avec les CHC.

CHC OPTIQUE 2D	
Paramètres du réseau	Valeurs
# Neurones entrées	6
# Couches cachées	1
# Neurones Couches cachées	28
# Neurones sorties	4
# Images entraînement	11
# Images tests	92
# Itérations	2380
Reconnaissance	99%
Discrimination	99%

On remarque tout d'abord le taux élevé de reconnaissance de l'ordre 99%. Le nombre d'itérations est de 2380. Ce nombre, tout à fait raisonnable pour un réseau de neurones (les poids étant initialisés aléatoirement), est tout de même nettement supérieur au 62 itérations requises pour l'apprentissage des CHCA. Néanmoins, les résultats sont encourageant. Un nombre plus élevé d'image tests permettrait de mieux caractériser les capacités et performances de cette méthode.

5.3 Commentaires généraux sur les CHC et CHCA

Les expériences réalisées ont permis de tirer des conclusions encourageantes sur l'utilisation des CHC et CHCA comme outils de classification. La reconnaissance a été démontrée de façon numérique malgré les translations, changements d'échelle et rotations dans le plan et hors-plan, alors que la reconnaissance complète en rotation et sous changement d'échelle a été démontrée à partir des données optiques. Les résultats obtenus lors des expériences optiques démontrent que le comportement des CHC et CHCA suit les mêmes tendances que les expériences réalisées numériquement. Ainsi, il est effectivement possible

d'extraire les CHC et CHCA de façon optique avec une stabilité suffisante pour réaliser des opérations de reconnaissance et classification.

Les réseaux de neurones complètent l'utilisation des CHC et CHCA en permettant la combinaison d'information recueillie sous différentes vues. Ceci est en grande partie dû à la possibilité de représenter l'image par un nombre réduit de CHC et CHCA, onze dans le cas présent. Les CHCA semblent démontrer une plus grande capacité de discrimination entre les différents objets que les CHC, ceci est reflété par le nombre réduit d'itérations requises pour l'entraînement des réseaux de neurones. Les CHCA ont aussi ouvert la possibilité d'extraire la rotation dans le plan d'un objet à partir de sa signature, d'autres tests et raffinements seraient souhaitables pour confirmer ce point. Il est à noter que les expériences optiques de rotation 3D des objets n'ont pas pu être réalisées faute de temps et de ressources. Ceci n'affecte pas cependant les conclusions pour les performances optiques des CHC et CHCA, puisque le cas 3D est une généralisation du cas bidimensionel.

Finalement, nous ajouterons que la stabilité des mesures optiques pourrait être améliorée par l'utilisation d'un seul ACL et de multiples rosettes encodées sous la forme de masques de chrome. Cette technique présenterait quelques avantages. L'utilisation d'un seul ACL permettrait d'éviter les battements rencontrés entre l'ACL des objets et celui des rosettes. Cette architecture se prêterait bien à un montage compact, les rosettes pouvant être apposées directement sur la surface du ACL. En considérant qu'un tel système pourrait traiter quatre rosettes par ACL, seulement quatre ACL seraient requis pour obtenir les 7 premiers ordres CHC et CHCA. De plus, l'utilisation d'un seul ACL augmenterait la quantité d'énergie lumineuse disponible, ouvrant la porte à l'utilisation d'une source à spectre étroit laquelle permettrait l'obtention d'un meilleur contraste sur l'ACL.

Références

[1] H. H. Arsenault et Y. Sheng, "Properties of the circular harmonic expansion for rotation-invariant pattern recognition", *Applied Optics*, Vol. 25, No. 18, 1986, pp. 3225-9.

[2] L. Neiberg et D. P. Casasent, "Feature Space Trajectory (FST) Classifier Neural Network", Proceeding SPIE, Vol. 2353, pp. 276-292, 1994

[3] C. T. Zahn et R. Z. Roskies, "Fourier descriptors for plane curves", IEEE Trans. on Computers, vol. c-21, 3, pp. 269-277, 1972.

[4] A. Mahalanobis et al., "A Quadratic Distance Classifier for Multi-class SAR ATR Using Correlation Filters", SPIE # 1875, pp. 84-95. 1993.

[5] J. W. Sherman et al., "Automatic Target Recognition Systems", IR/EO Handbook.

[6] A. Bergeron et al., "Programme de recherche sur les neurones optiques appliqués à la reconnaissance de cibles", Rapport de l'Institut national d'optique INO 94-7404-2RC RET N/A (Rapport de mi-projet).

[7] J. W. Goodman, "Introduction to Fourier Optics", McGraw-Hill Book Company, San Francisco, 1968.

[8] R. N. Bracewell, "The Fourier Transform and Its Applications, Second Edition", McGraw-Hill Publishing Company, New York, 1986.

[9] M. R. Spiegel, "Formules et tables de mathématiques", McGraw-Hill, New York, 1985.

[10] L. E. Scales, "Introduction to Non-Linear Optimization", Springer-Verlag, New York, 1985.

[11] G. Bruhat, "Cours de Physique générale, Optique, sixième édition", Masson, Paris, 1992.

ANNEXE A

Assemblage du systèmes à réseau de diffraction

L'assemblage du système est une opération délicate. Chacune des rosettes doit être positionnée avec une erreur de centrage de l'ordre de quelques µm relativement aux répliques de l'image d'entrée. Les prochains paragraphes portent sur les aspects directement liés à l'assemblage optique du prototype développé à l'INO.

A.1 Mécanique

Pour permettre le positionnement individuel des rosettes avec une grande précision en tenant compte du peu d'espace disponible, il a fallu concevoir et construire un dispositif mécanique de précision. Puisqu'il était très difficile de faire un mécanisme de précision qui puisse se loger dans les espaces entre les rosettes, le mécanisme construit est amovible. Il permet d'ajuster une seule rosette à la fois. Lorsque le positionnement est complété, la rosette est immobilisée à l'aide de deux petites vis. La figure A. 1 montre le schéma du dispositif amovible de positionnement et du support de fixation des rosettes.

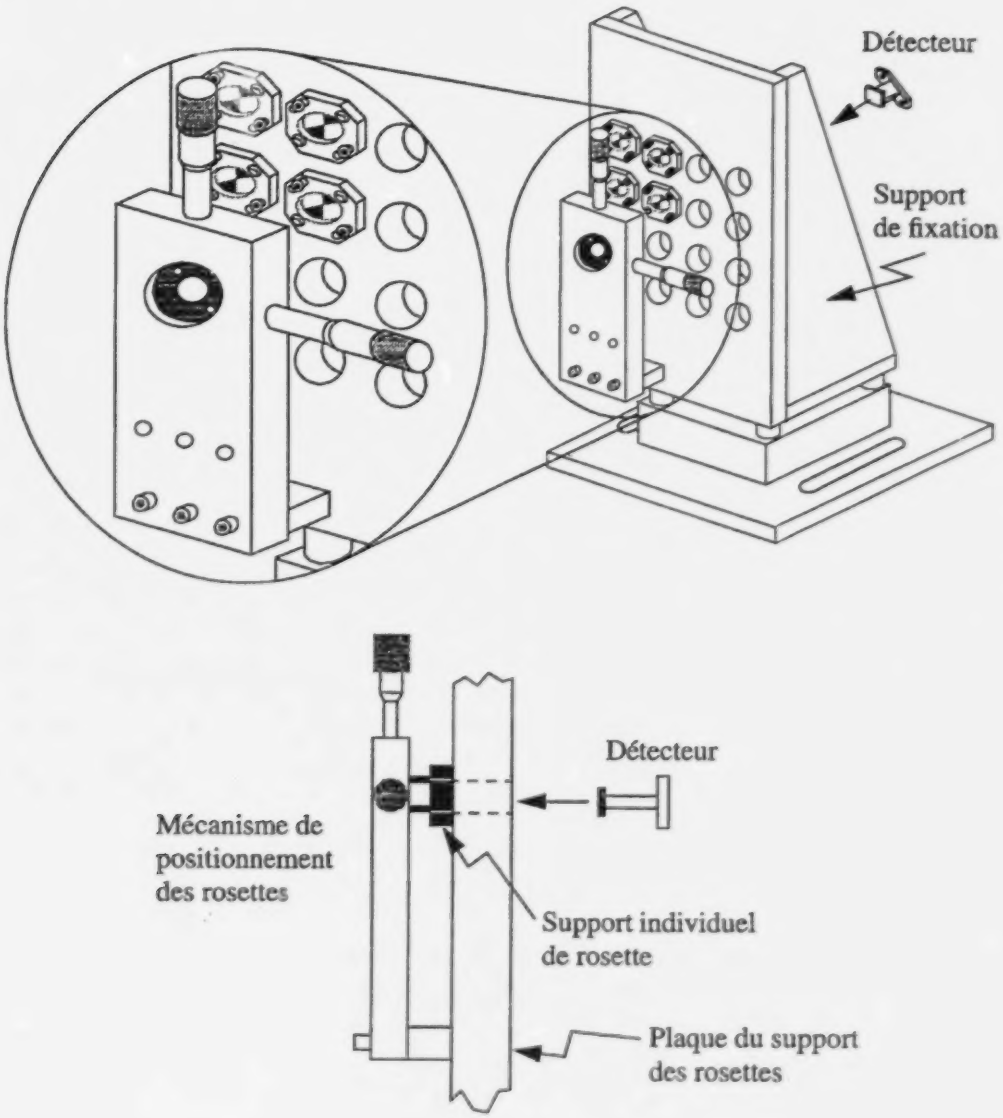


FIGURE A. 1: Mécanisme de positionnement et dispositif de fixation des rosettes vue en perspective en haut. Vue de profil du mécanisme de positionnement en bas.

Comme le montre la figure A. 1, les rosettes sont montées dans des supports individuels. Le support de fixation qui assure l'immobilisation des rosettes sur la table optique comprend une plaque plane dotée de grands orifices qui sont destinés à accueillir les détecteurs. Les rosettes sont fixées sur la surface avant de la plaque vis-à-vis des orifices tandis que les détecteurs sont introduits dans ceux-ci et fixés sur la surface arrière de la plaque. Le mécanisme de positionnement comprend deux tiges dont les extrémités sont dotées d'un mécanisme de ressort. Ces tiges permettent d'agripper les supports individuels des rosettes par l'entremise de 2 concavités. Le mécanisme de positionnement est boulonné sur le support de fixation et les ressorts des tiges exercent une pression qui maintient la rosette bien appuyée contre la surface avant du support de fixation. Les tiges du mécanisme de positionnement sont montées sur une bague cylindrique, laquelle est imbriquée dans une bague plus grande. Le tout peut être déplacé verticalement et horizontalement à l'aide de deux vis micrométriques et une autre vis permet de faire tourner la bague centrale. La plage de déplacement est de ± 1 mm selon les deux axes tandis que la rotation peut se faire sur ± 5^0.

La figure A. 7 montre une vue agrandie d'une rosette dans son support individuel. On y voit les concavités par où s'agrippe le mécanisme de positionnement. On y voit également les 2 orifices surdimentionnés par où se glisse les vis d'immobilisation.

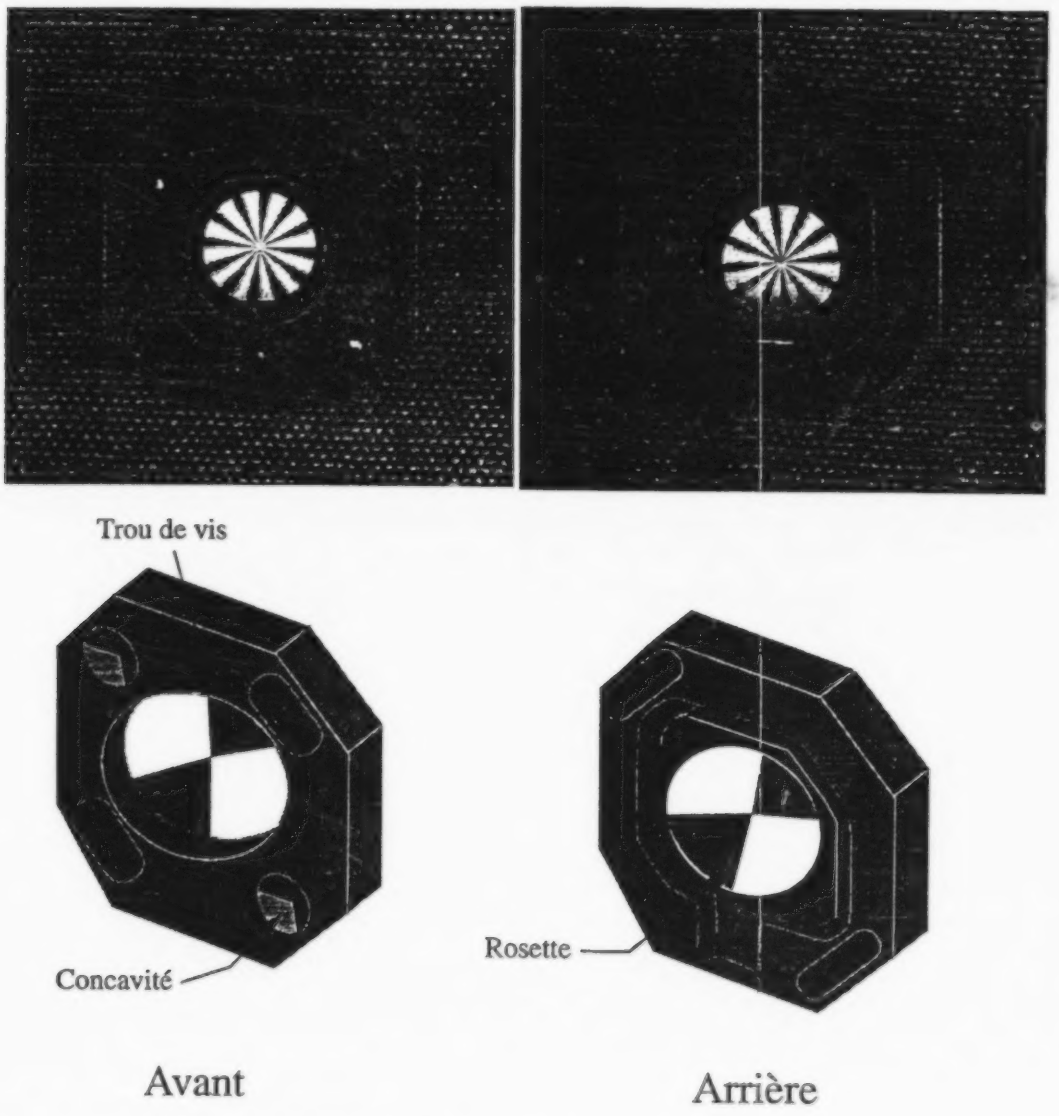
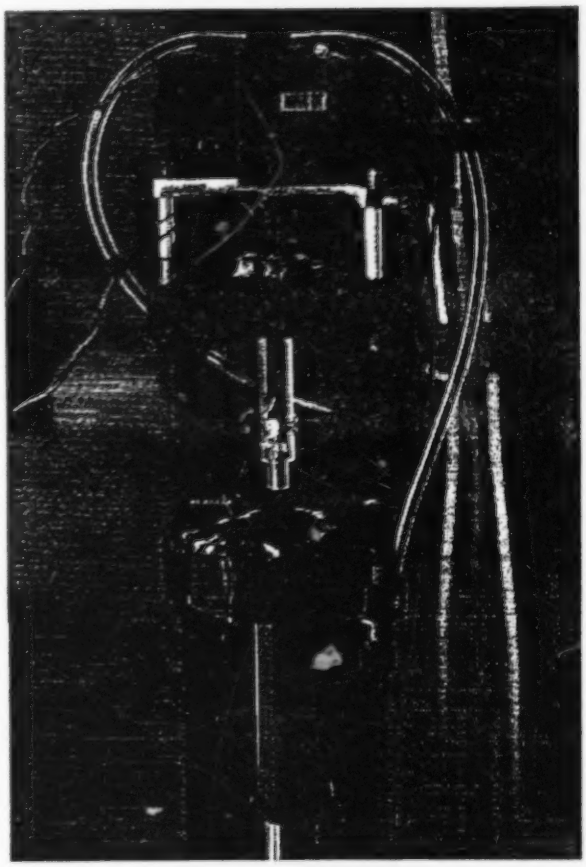


FIGURE A. 2: Support individuel des rosettes

Étant donné la faible marge de manoeuvre offerte par le mécanisme de positionnement (± 1 mm, ± 5°), les rosettes doivent être collées dans leurs supports individuels avec précision. Le collage des rosettes dans leurs supports a été réalisé à l'aide du montage illustré à la figure

A. 3. Dans ce montage les rosettes sont positionnées dans le centre du support par l'entremise d'un tige transparente montée sur un mécanisme de translation et de rotation. Les rosettes sont maintenues en place au bout de la tige transparente par succion. Le suivi se fait à l'aide d'une caméra vidéo équipée d'une optique permettant un grandissement d'environ 6-7 X, ce qui est suffisant pour voir les pixels des rosettes. L'éclairement est assuré par une DEL et se fait au travers de la tige transparente. La caméra est montée sur une table de translation à trois axes permettant de faire la mise au point et de la positionner au centre du support individuel de la rosette. La rosette est déplacée de façon à ce que son centre soit confondu avec l'emplacement du champ de la caméra, préalablement identifié comme étant le centre du support de rosette. L'orientation de la rosette est ajusté en même temps que se fait son centrage. La rosette est tournée sur elle même jusqu'à ce que ses pixels soient orientés selon une ligne de référence placée sur le moniteur de la caméra. Ceci nous assure que les rosettes soient orientées de façon semblable relativement à leur support respectif. La fixation des supports de rosettes sont fixés selon une orientation bien précise relativement à la caméra. Une fois le positionnement complété, la rosette est collée définitivement sur son support à l'aide d'une colle polymérisable par exposition au rayonnement ultraviolet (Norland Optical Adhesive).



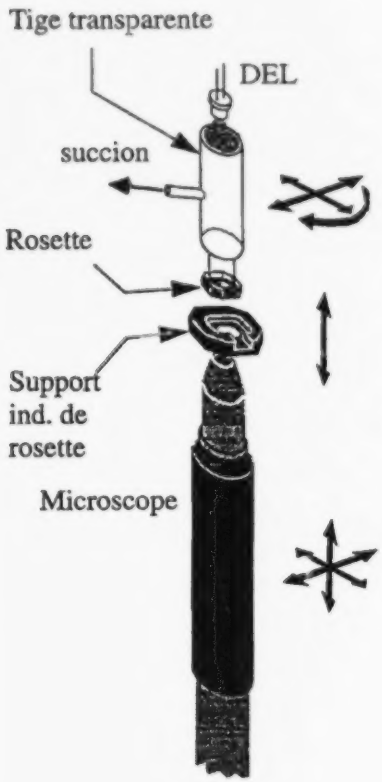


FIGURE A. 3: Montage pour le collage des rosettes

La précision du positionnement des rosettes dans leurs supports individuels a été vérifiée à l'aide d'un microscope de précision. Il a été estimé que les erreurs sur le centrage sont inférieures à 200 µm tandis que celles sur l'orientation sont inférieures à 0.5^0. Ces résultats sont très satisfaisant puisque les erreurs sont bien en deçà des valeurs des plages accessibles avec le mécanisme de positionnement des rosettes (± 1 mm, ± 5^0).

Le positionnement des filtres spatiaux demande beaucoup moins de précision que celui des rosettes. Aussi, le mécanisme de positionnement des filtres spatiaux est beaucoup plus simple que celui des rosettes. Celui-ci est illustré à la figure A. 4

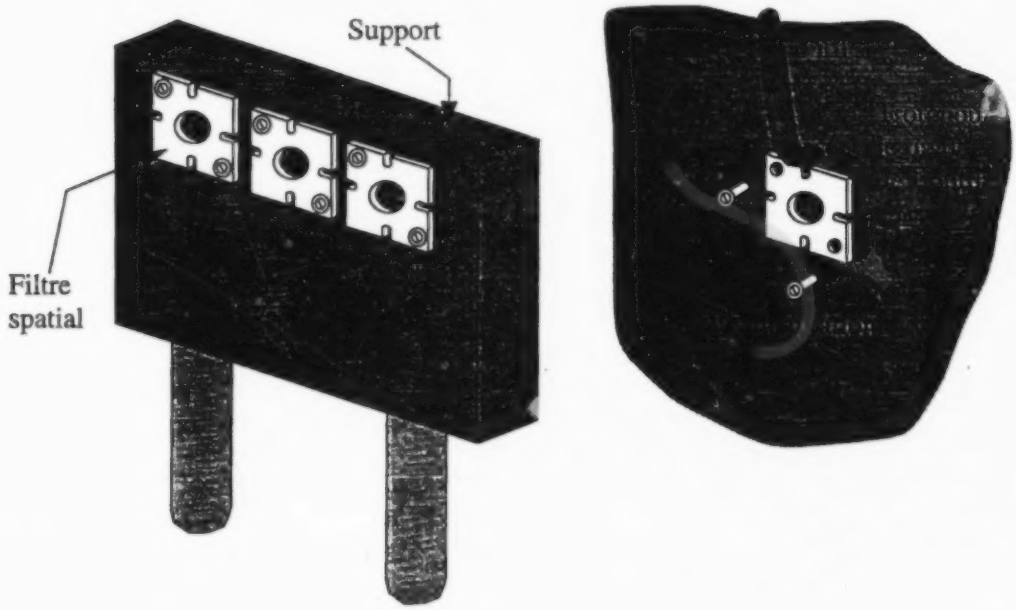


FIGURE A. 4: Filtres spatiaux, leur support et l'outil de déplacement.

Le support est une simple plaque plane doté de grands orifices. Les filtres spatiaux sont de petite plaques percées en leur centre d'un trou de diamètre 4 mm. Les filtres comprennent également 4 petites fentes dans lesquelles les outils de déplacement prennent appui. Un de ces outils est illustré à la droite de la figure A. 4 avec un filtre spatial. La plaque (support) est dotée de plusieurs petits trous disposés autour des filtres spatiaux permettant l'utilisation de quatre outils simultanément sur un même filtre.

Le positionnement des filtres spatiaux demande beaucoup moins de précision que celui des rosettes. Aussi, le mécanisme de positionnement des filtres spatiaux est beaucoup plus simple que celui des rosettes. Celui-ci est illustré à la figure A. 4

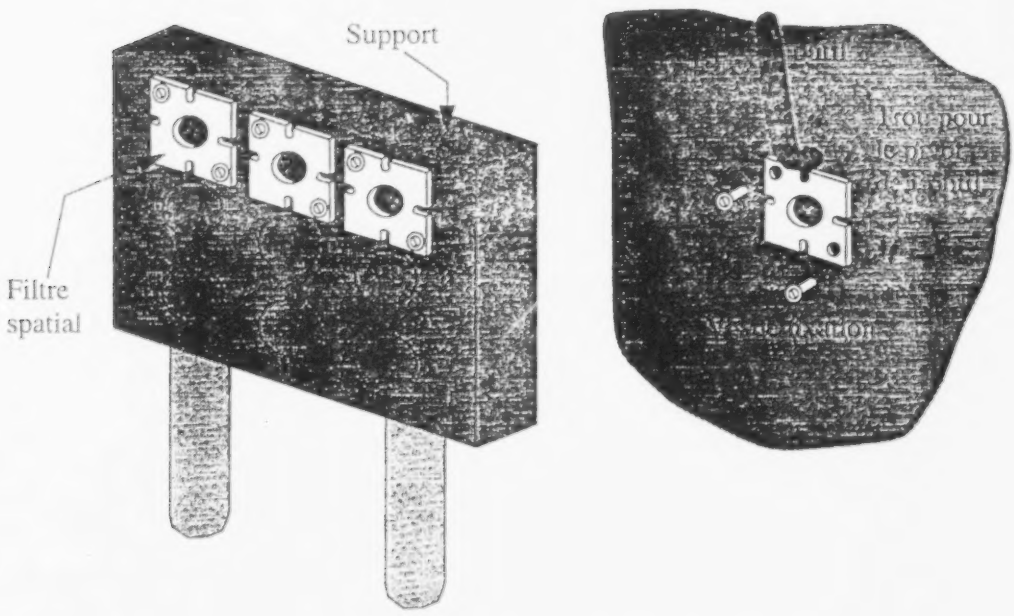


FIGURE A. 4: Filtres spatiaux, leur support et l'outil de déplacement.

Le support est une simple plaque plane doté de grands orifices. Les filtres spatiaux sont de petite plaques percées en leur centre d'un trou de diamètre 4 mm. Les filtres comprennent également 4 petites fentes dans lesquelles les outils de déplacement prennent appui. Un de ces outils est illustré à la droite de la figure A. 4 avec un filtre spatial. La plaque (support) est dotée de plusieurs petits trous disposés autour des filtres spatiaux permettant l'utilisation de quatre outils simultanément sur un même filtre.

A.2 Alignement

L'alignement du système est une opération relativement complexe en raison du nombre important de composants impliqués et de la précision élevée requise. L'alignement se fait en plusieurs étapes et débute par un préalignement destiné à établir les positions approximatives des composants sur la table optique. Les prochains paragraphes décrivent sommairement les étapes plus critiques.

Étant donné la marge de manoeuvre limitée au niveau du positionnement des rosettes, le support de fixation des rosettes doit constituer l'élément de référence, celui-ci dicte l'ajustement des autres composants. L'axe optique du système doit être perpendiculaire à sa surface en plus d'être centré sur l'emplacement nominal prévu pour la rosette centrale. Ceci implique que le faisceau laser (pas encore dilaté) doit être orienté (par réflexion sur 2 miroirs ajustables) de façon à répondre à cette condition.

L'espacement entre les répliques des images doit être très proche de la valeur de l'espacement nominal entre les rosettes sur le support de fixation. Or comme le montre la figure A. 5, l'espacement entre les répliques dépend de la distance qui sépare la dernière lentille (Lentille Oriel 750 mm) du plan des rosettes. Le plan des rosettes est située approximativement dans le plan focal de la lentille ($d \cong f$). L'espacement E entre les répliques des images est approximativement donné par $E = d\theta$, θ étant l'angle de déviation du premier ordre du réseau ($\theta \cong \lambda/p$, p étant le pas du réseau c'est-à-dire la dimension de la cellule élémentaire). Pour s'assurer que les centres des répliques seront à l'intérieur des marges de déplacement des rosettes il faut déplacer la lentille de façon à ce que E corresponde à la distance nominale prévue. Cet ajustement se fait lorsque les seules composants présents sur la table sont le support de fixation des rosettes, le réseau de diffraction et la dernière lentille. L'ajustement de l'orientation (rotation autour de l'axe optique) du réseau de diffraction doit se faire en même temps. Le réseau doit être tourné de façon à ce que ses axes soient alignés selon les directions de l'agencement des rosettes.

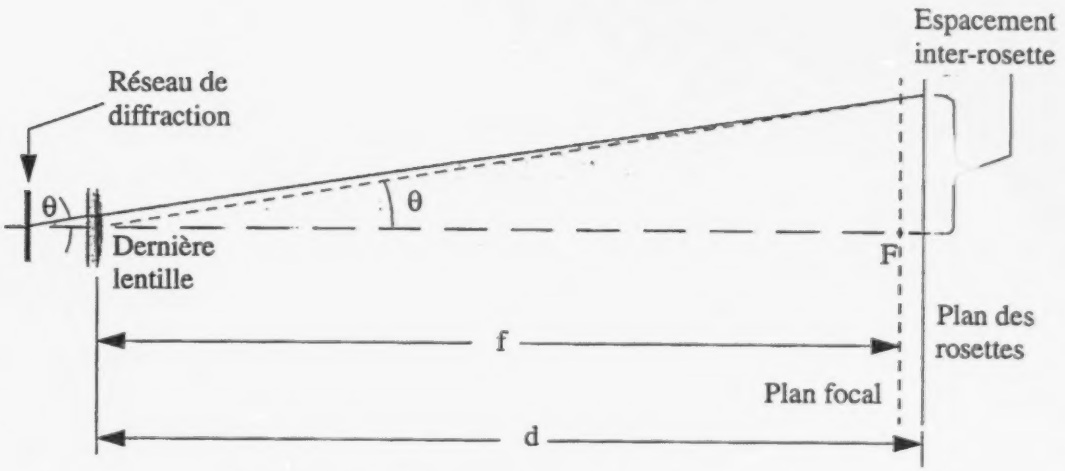


FIGURE A. 5: Espacement entre les ordres de diffraction

Après avoir dilaté le faisceau laser et avoir ajusté l'ensemble des lentilles, il faut centrer les rosettes. Ceci se fait avec un réticule centré dans le plan d'entrée du système à l'endroit même prévu pour l'ACL. Les répliques de l'image du réticule sont projetées sur les rosettes. À l'aide du mécanisme de positionnement et d'une caméra vidéo dotée d'une optique de microscope, les centres des rosettes sont amenés à tour de rôle en coïncidence avec le centre du réticule. Lors du centrage, les rosettes sont également orientées selon une direction commune grâce au réticule. La caméra vidéo est placée à l'arrière du support de fixation des rosettes et l'observation se fait par les trous prévus pour accueillir les détecteurs. La procédure d'ajustement des rosettes est illustrée à la figure A. 6.

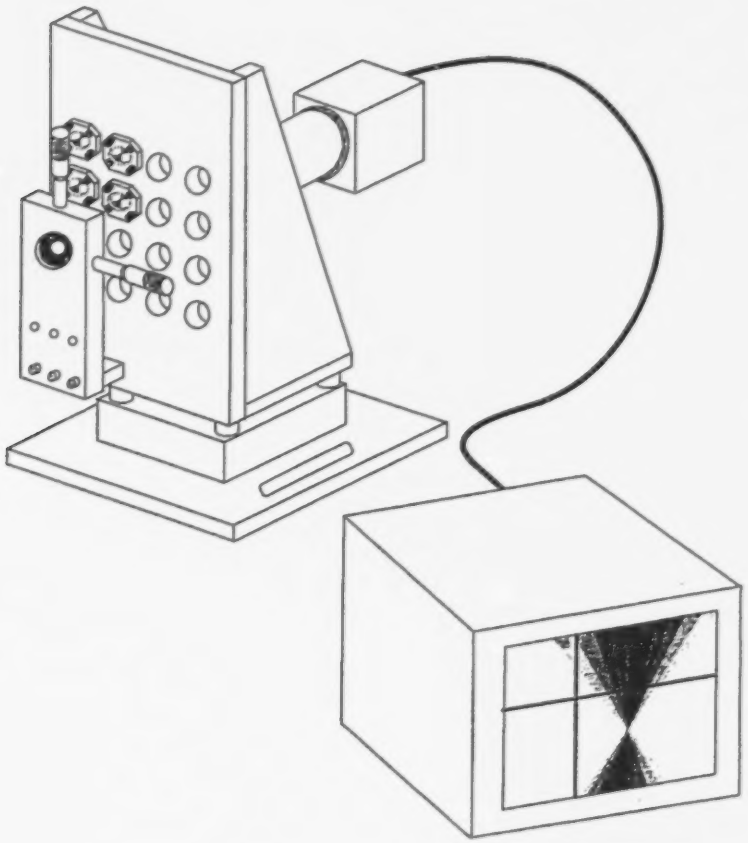


FIGURE A. 6: Ajustement des rosettes.

À la fin du processus d'alignement, lorsque l'ACL a été mis en place il faut faire l'ajustement des filtres spatiaux. Ceci se fait en utilisant une caméra vidéo ayant une optique de microscope. L'observation en cours d'ajustement se fait par l'arrière. La transformée de Fourier et ses répliques générées par la pixellisation sont observées à l'aide d'une pellicule diffusante collée sur la face avant du filtre spatial.

Lorsque tous les ajustements sont complétés il ne reste qu'à insérer les détecteurs dans leurs orifices derrière les rosettes. Un centrage grossier suffit pour les détecteurs puisque ceux-ci sont légèrement plus grand que les rosettes et qu'ils sont placés très près des rosettes. La figure A. 7 présente des photographies du montage une fois tous les ajustements complétés.

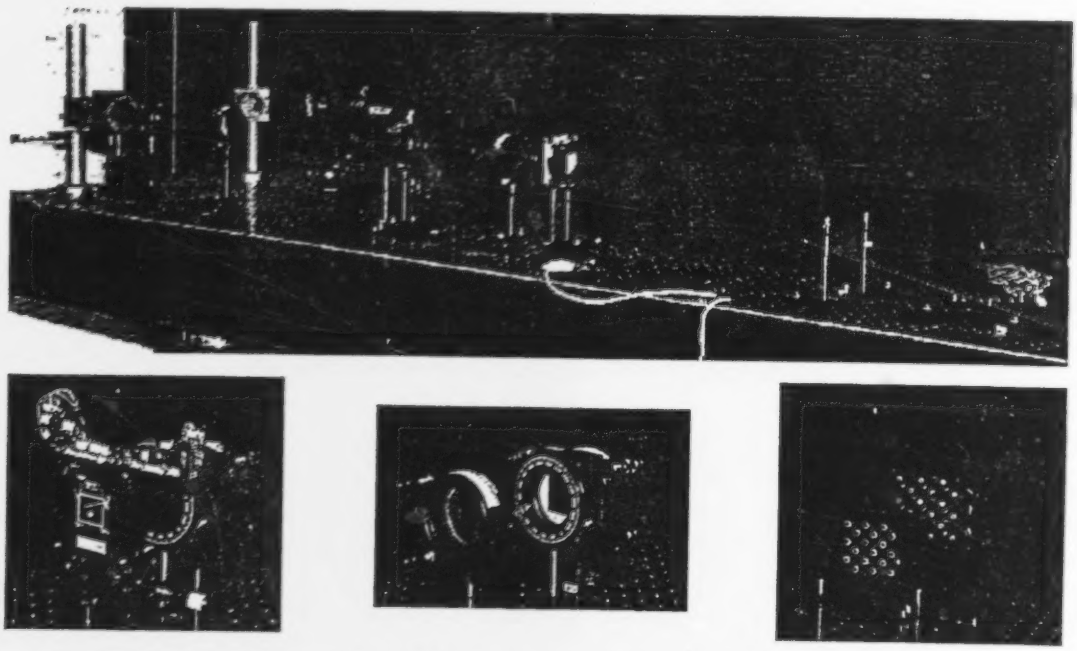


FIGURE A. 7: Photographies du montage optique complété avec gros plans en bas à droite sur l'ACL, en bas au centre sur le réseau de diffraction et en bas à gauche sur les filtres spatiaux et les rosettes.

Seconde partie: Contribution du COPL

Introduction

6. Introduction

Ce contrat porte sur la reconnaissance des formes invariante sous translation, rotation en 3-D et sous changement d'échelle. Cette étude porte donc sur les techniques permettant de faire de la classification de cibles (objet tridimensionnels) terrestres indépendamment de l'orientation, de la position et de l'échelle des cibles, les cibles étant perçues a partir du sol par une caméra infrarouge.

À première vue, on pourrait penser que pour obtenir simultanément l'invariance sous rotation et sous échelle, il suffit de combiner une méthode invariante sous rotation avec une méthode invariante sous échelle; malheureusement il n'en n'est rien. Une méthode invariante sous rotation n'est invariante que pour une seule échelle, et une méthode invariante sous échelle n'est valable que pour une seule orientation. Il est donc clair que si l'orientation de la cible et son échelle changent tous les deux, aucune des deux méthodes ne fonctionnera.

Comment donc faire pour obtenir les deux invariances simultanément? Une approche possible consiste à modifier l'image pour qu'elle soit invariante sous changement d'échelle ou sous changement d'orientation, et de lui appliquer ensuite une méthode qui satisfasse à l'autre invariance. Une autre approche consiste à extraire des paramètres invariants de la cible, et de classer les images à partir de ces invariants. L'une des méthodes que nous avons choisies utilise la première approche, alors que l'autre utilise la seconde.

Il y a deux classes de problèmes dont il convient d'être conscient dans une étude comme celle-ci. La première est liée au problème théorique lui-même: on a une ou plusieurs classes d'objets idéaux qu'on veut classer, et on cherche le moyen de le faire. Le second problème est lié aux sortes d'objets considérés et à leur représentation.

Pour ce contrat, les objets à reconnaître sont des images vidéos infrarouges de véhicules militaires. De telles images sont peu contrastées et entachées de bruit. Donc si une méthode ne fonctionne pas bien, il faut se poser la question si cela est dû à la mauvaise qualité des images ou à la méthode elle-même.

Idéalement, la méthode scientifique exigerait qu'on étudie les méthodes proposées d'abord avec des objets bien définis et sans bruit, pour ensuite passer à des images plus réalistes: cela permettrait de déterminer si certaines difficultés rencontrées sont intrinsèques à la méthode ou causées par la mauvaise qualité des images.

Au cours de cette étude, nous avons pu constater que certains effets causés par la numérisation des images sont dominants et sont la principale cause d'erreurs dans le classement. L'absence d'une banque d'images ayant des échelles variées s'est avérée une lacune importante dont nous discuterons plus loin.

Bien que l'une des méthodes retenues puisse se réaliser optiquement, ce contrat ne comporte pas de partie optique. Les résultats obtenus peuvent donc être considérés comme une simulation de l'expérience optique, ou comme une expérience purement numérique.

6.1. Tâches à accomplir

Les tâches à accomplir dans le cadre de ce contrat sont les suivantes:

1. Faire une recherche sur les techniques optiques, optoélectroniques et neuronales susceptibles de contribuer a la réalisation des objectifs du projet, identifiées comme devant être accomplies par l'Université Laval dans le cadre du projet confié à l'INO par le CRDV (proposition INO 93-7404-2 P N/A).

   Cette recherche consiste à:

   - passer en revue les principales techniques
   - décrire sommairement les techniques
   - donner les avantages et les désavantages de ces techniques
   - faire une liste des références pour chacune des techniques.

2. Appliquer deux des techniques les plus prometteuses choisies de concert avec l'INO pour la reconnaissance invariante selon la position, l'échelle et l'orientation (rotation hors-plan et dans le plan). Les tests seront effectués au moyen de simulations par ordinateur sur des images fournies par l'INO/CRDV contenant des objets d'intérêt.

3. Déterminer le rendement de chacune des méthodes considérées dans des conditions différentes plus ou moins complexes (addition de bruit, de fond de terrain, d'objets parasites, de changement de contraste).

4. Rédiger des rapports sur les résultats obtenus.

5. Ajouter aux rapports un sommaire des résultats des projets en cours dans les laboratoires des chercheurs de l'Université Laval sur la reconnaissance invariante et les réseaux neuroniques optiques qui ne sont pas couverts par ce contrat.

6.2. Personnel

Le personnel qui travaille sur ce contrat consiste en deux professeurs du département de physiques et de deux étudiants de 2ième et de 3ième cycles. Ce rapport a été rédigé par les quatre participants. Une partie du rapport est extraite de la thèse de Maîtrise de Sébastien Parent.

Professeurs: Henri H. Arsenault (responsable scientifique)

Yunlong Sheng

Etudiants: Sylvain Deschênes

Sébastien Parent

Autorité scientifique à l'INO: Michel Doucet

6.3. Revue des principales techniques

Cette section présente une bibliographie commentée des méthodes d'intérêt pour ce contrat. Disons tout de suite que sauf quelques exceptions, cette étude exhaustive des écrits a livré peu de résultats. Au moment où ce contrat a été émis, il n'existait aucune méthode satisfaisante de reconnaissance des formes ayant subi des rotations en 3-D.

Beaucoup de travaux sont faits sur les reconnaissances invariantes en 2-D comparativement à celles en reconnaissance 3-D. Quelques unes sont numériques (réseau de neurones[1,2,3], méthode des moments[4], coincidence des points d'inflexion du contour d'objet[5] et autres méthodes[6,7,8,9]). D'autres sont implémentables optiquement (réseau de neurones[10,11], filtres d'harmonique circulaire[12,13,14], "synthetic discriminent filter[15]", GRI-MACE[16]). Ces méthodes ne sont cependant pas étudiées et optimisées pour la reconnaissance sous distortion comme il se produit dans la cas de la rotation hors plan. C'est pourquoi elles ne sont pas retenues comme étant des méthodes utilisables pour la reconnaissance 3-D.

Une question importante qui se pose tout d'abord, étant donné la complexité du problème abordé, est de savoir si l'objectif visé est réalisable. En effet, la reconnaissance seulement des formes invariantes sous des rotations arbitraires en 3-D est un problème non-résolu, et notre but ici est la reconnaissance invariante non seulement sous de telles rotations, mais également invariantes sous des changements d'échelle et de translation!

Or des chercheurs ont démontré[17] qu'en effet, *il n'y a pas d'invariants pour des rotations arbitraires en trois dimensions.* Il convient d'examiner cette démonstration, afin de déterminer les limitations qu'elle pose pour ce travail. À l'aide d'une série de théorèmes et de lemmes reliés aux transformations affines (opérations de transformations linéaires en 3-D et de translation) et aux diverses projections de l'objet 3-D, les auteurs de cette démonstration montrent que l'unique fonction invariante sous ces transformations est une constante et donc ne dépend pas de l'objet. Cette solution étant triviale il devient clair qu'il n'existe aucun façon de trouver un invariant pour toutes les classes d'objets 3-D considérées sous toutes les perspectives.

Dans les prochaines sous-sections, nous verrons les méthodes de reconnaissance 3-D qui existent mais qui ne sont pas retenues pour le contrat et nous ferons une brève introduction des deux méthodes qui ont été choisies et qui font l'objet d'études plus approfondies.

6.4. Techniques non retenues

À la suite de la preuve de l'inexistance d'un invariant pour la rotation hors plan sur un intervalle de 360 degrés, certains auteurs ont donc décidé d'analyser le problème d'une façon moins générale en étudiant des groupes spécifiques d'objets ou de transformations. Il sont ainsi en mesure de trouver un invariant pour cette classe d'images particulière. Ainsi, certains[18] trouvent une fonction invariante pour les objets de faible perpective et dérivent un modèle qui permet de relier les invariants de chacun des groupes de perspectives en une seule fonction caractérisant les objets pour toutes les perspectives. D'autres[19] utilisent des descripteurs invariants pour une classe particulière d'objets, soit les objets ovoïdes plans. Les descripteurs invariants sont reliés à la forme des objets et peuvent reconnaître toutes les projections et toutes les transformations linéaires pour la classe d'objet utilisée.

Le désavantage majeur des deux méthodes décrites précédemment est qu'elles nécessitent une représentation des objets par des points fixes relevés sur ces objets. Ainsi, il faut être en mesure de trouver la positions exacte des points d'analyses sur les objets, sans connaître l'orientation de l'objet dans l'espace, avant de pouvoir faire les analyses sur l'objet.

Ullman et Rasri ont développé[20] un modèle mathématique permettant de reconnaître une image inconnue à partir d'une combinaison linéaire d'images connues ayant subit des transformations affines. Cependant, cette méthode n'est pas développée expérimentalement.

Au niveau des méthodes testées expérimentalement, le système SAHTIRN (Self Adaptive Hierachical Target Identifiction and Recognition Neural Network), construit par le groupe de Cindy Daniell de Hughes Aircraft, porte sur la détection de véhicules sur le terrain à partir d'images infrarouges[21]. Après un rehaussement du contraste et une extraction de contours, un réseau de neurone Néocognitron[22] extrait les caractéristiques primitives des objets. Finalement, un réseau de rétropropagation procède à la classification. Le SAHTIRN réussit à atteindre une classification correcte à 99% pour le cas 2D sur une base de données limitée. Le Néocognitron procède à un apprentissage non-supervisé qui ajuste automatiquement les synapses entre les couches. Il reconnaît des images invariantes sous translation et avec une certaine tolérance à la distorsion des images.

Un autre système traitant le même type de problème est le XTRS par Verly et al. du MIT, Lincoln Lab.[23] Il est basé sur un "matching" entre les contours des images et un système hiérarchique muni d'une librairie de primitives servant à décomposer le contour trouvé. Le modèle de description structurale est basé sur les parties de l'objet et leurs interrelations spatiales. Les parties et les relations sont codées par les neurones dans différentes couches du réseau. Le réseau dérive une description structurale de l'objet en fonction de ses parties et de ses interrelations[24]. Ce modèle est cependant sensible à la rotation dans le plan de l'image. Le désavantage de cette approche réside dans la lenteur du calcul pour regrouper les relations.

Sans entrer dans les détails, nous pouvons également mentionner l'existence d'un autre système, le MODALS, proposé par Thoet et al. de la cie Booz-Allen and Hamilton.[25] Disons seulement que cette technique est basée sur la mémoire associative.

Une autre méthode[26] porte sur l'utilisation d'un réseau de neurones à sorties pulsées. Ce système donne un signal périodique qui est utilisé comme signature pour caractériser les objets. Il est démontré par simulation numérique que les signatures ont une invariance en rotation sur une plage de 30 degrés, une invariance sous échelle pour un facteur 2 et une invariance à la distorsion équivalente à une rotation hors plan de 30 degrés. Cependant cette méthode est testée avec des objets simples (lettre T et symbole +) et avec des objets plans.

6.5. Techniques retenues

Deux méthodes ont été retenues. Ce sont celles qui seront étudiées plus en détail dans le cadre du projet. La première utilise la compression d'information pour permettre des invariances en translation et sous échelle. Une banque de données caractéristiques est ensuite construite en mémoire pour chacun des objets à reconnaître. La seconde technique utilise l'extraction des contours et analyse ces contours avec des descripteurs de Fourier pour avoir les invariances sous translation, sous échelle et sous rotation dans le plan. Ces deux méthodes ont l'avantage d'avoir une bonne robustesse face à la distortion des images. Les deux prochains chapitres portent chacun sur le fonctionnement et sur les résultats obtenus pour chacune des techniques.

6.6. Description de la banque d'images utilisée

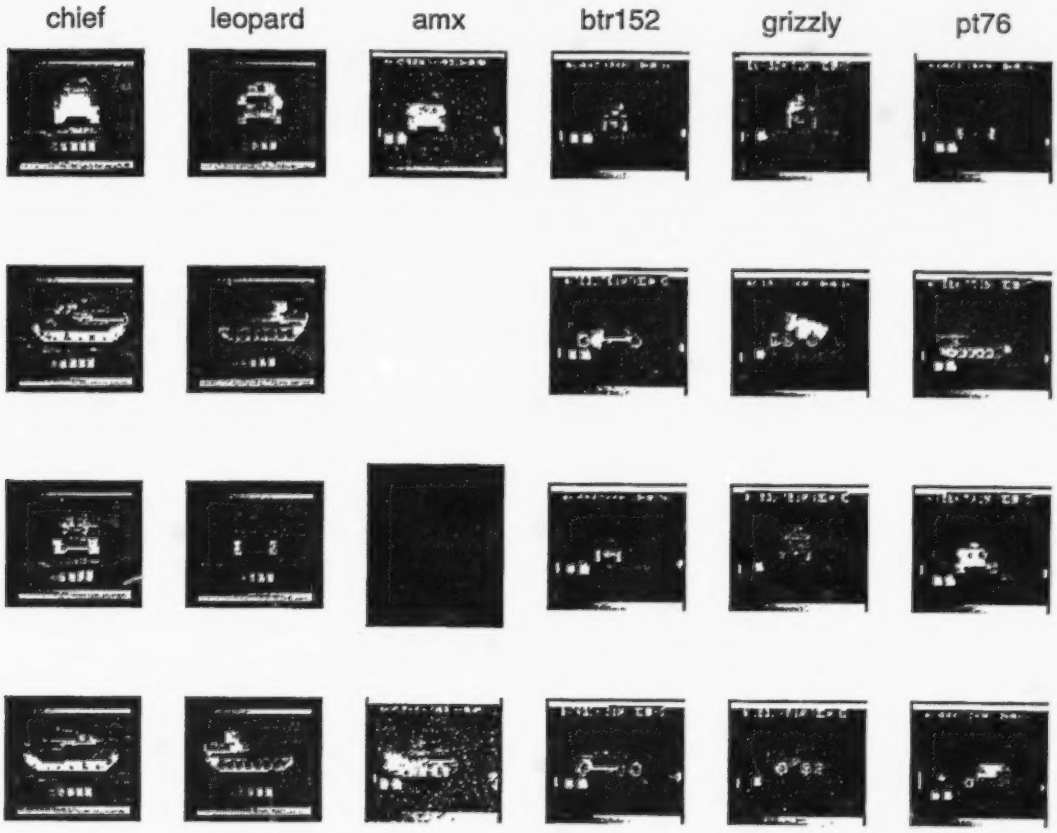


Figure 1: Banque d'images utilisée pour le projet

La banque d'images utilisée contient six objets différents. Ce sont tous des véhicules militaires que l'on divise en deux catégories. La première est la classe des chars d'assaut: le chief, le léopard et le amx. La seconde est celle des transporteurs de troupe: le btr152, le grizzly et le pt76. Les images de chacun des objets sont prises à différents points de vue. L'écart de rotation entre chaque point de vue est de 5 degrés. Cinq des six véhicules comporte une série de 72 images couvrant toute la région de la rotation hors plan, soit de 0 à 360 degrés. Seul le amx n'a pas toute la plage de rotation. De plus, les images de ce dernier contiennent beaucoup de bruit. Pour ces deux raisons, nous mettrons de côté l'analyse de ce véhicule et nous porterons notre étude sur les cinq autres objets.

Les images ont été captées par le CRDV à l'aide d'une caméra infrarouge. Elles ont des dimensions de 512 X 480 pixels et sont compressées en format « raster file ». En regardant les images initiales, on remarque un rectangle sombre sur lequel repose les véhicules. C'est le plateau utilisé pour faire tourner les véhicules en rotation hors plan. De plus, on remarque un peu plus bas une suite de carrés d'intensités différentes. C'est la charte d'intensité de la caméra. Ces deux dernières informations ne sont pas utiles lors du traitement des images et elles seront éliminées lors de la segmentation.

Première méthode

7. Description de la méthode des trajectoires d'espace caractéristique avec réseau de neurones (FST NN)

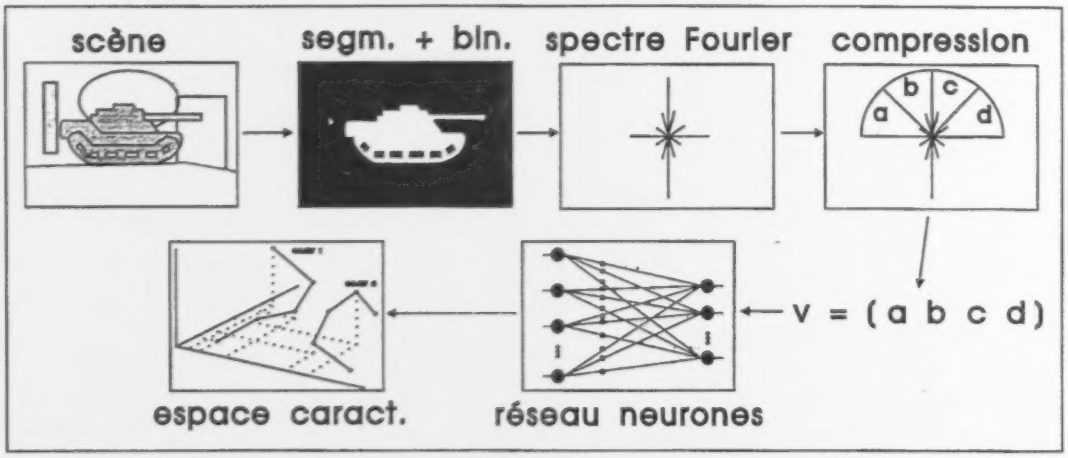


Figure 2: Schéma du système utilisé par Neiberg et Casasent

Une des méthodes retenues a été développée par Leonard Neiberg et David P. Casasent[27]. Cette méthode consiste à compresser l'information contenue dans une image et à faire l'apprentissage et la reconnaissance des objets à partir de l'information des images compressées. Cette méthode est appelée par les auteurs le « feature space trajectory classifier neural network FST NN ». La Figure 2 montre un schéma de ce système. Nous détaillerons ici le fonctionnement de chacune des étapes du traitement tel qu'utilisé par les auteurs de la méthode et nous verrons dans le chapitre suivant que quelques modifications à la méthode ont été apportées dans le cadre de notre projet. Les détails de ces changements y seront expliqués.

7.1. Segmentation

La première opération que subit une image entré dans le système FST NN est un pré-traitement qui consiste à faire la segmentation des objets d'intérêt dans une scène quelconque. Cette opération permet d'isoler les objets à étudier et de les placer dans une fenêtre de dimension fixe. Ces images sont ensuite binarisées. La banque d'images ainsi prétraitée contient six différents objets avec des rotations hors plan variant entre -90 et 90 degrés. Chaque objet possède 21 images de rotation différentes. L'apprentissage se fait sur 11 de ces images et la reconnaissance sur les 10 autres images. L'apprentissage s'effectue à environ tous les 15 degrés de rotation des objets et l'écart minimum entre deux images d'un même objet est de 5 degrés.

7.2. Spectre de puissance

Une fois le pré-traitement effectué, on fait la transformée de Fourier des images et on prend leur spectre de puissance. L'intérêt d'une telle opération est l'invariance sous translation qu'elle apporte aux images. Ainsi, peu importe où se situe l'objet dans la fenêtre d'observation, il sera repéré et reconnu.

7.3. Compression d'image

L'étape suivante du système consiste à réduire la dimension de l'information traitée. Pour ce faire, nous utilisons un système qui intègre l'énergie de régions déterminées du spectre de Fourier en une seule valeur par région. La sortie du système à cet endroit est donc un vecteur de même dimension que le nombre de secteurs du compresseur. En donnant une structure en secteurs (pointes de tarte) à ces régions, nous obtenons une invariance sous échelle. On appelle ce compresseur un « wedge Fourier filter[28,29] ». Avec le spectre de Fourier et avec ce type de compresseur, le système permet donc d'avoir l'invariance sous translation et l'invariance sous échelle. Il nous manque l'invariance sous rotation dans le plan, mais elle est négligée par l'hypothèse que les angles de dépression des objets utilisés (véhicules) est faible. On suppose ainsi que les objets à reconnaître se promènent sur un terrain plat ou faiblement accidenté.

7.4. Réseau de neurones

Une fois les images compressées en vecteurs, on passe cette information à travers un réseau de neurones. Le réseau utilisé ne comporte qu'une couche de neurones d'entrée et une couche de neurones de sortie. Il ne contient aucune couche cachée et aucune non linéarité. Ainsi, si on observe un neurone de sortie élémentaire de ce système on trouve l'équation suivante:

$$y_j = \sum_i w_{ij} \cdot x_i$$

où w_{ij} est le poids correspondant au neurone d'entrée i et de sortie j, x_i est le neurone d'entrée i et y_j est le neurone de sortie j. Nous utilisons comme poids w la banque d'image de l'apprentissage. Ainsi, si l'apprentissage comporte 5 images, il y aura 5 neurones de sorties correspondants à la corrélation entre le vecteur d'entrée et les vecteurs d'apprentissage.

7.5. Trajectoire d'espace caractéristique

Une fois sortie du réseau de neurones, le nouveau vecteur a la même dimension que le nombre d'images apprises (N_A). Chacun des vecteurs peut être représenté par un point dans un espace de dimension N_A appelé espace caractéristique. Pour représenter chacun des objets à reconnaître, on place dans l'espace caractéristique les vecteurs de sortie correspondants aux images d'apprentissage de chacun des objets et on relie ensuite ces points caractéristiques par des droites. On se retrouve avec une courbe en segment de droite appelée trajectoire d'espace caractéristique. L'ordre dans lequel on relie les vecteurs est important. C'est le même ordre que celui suivi lors de la rotation des

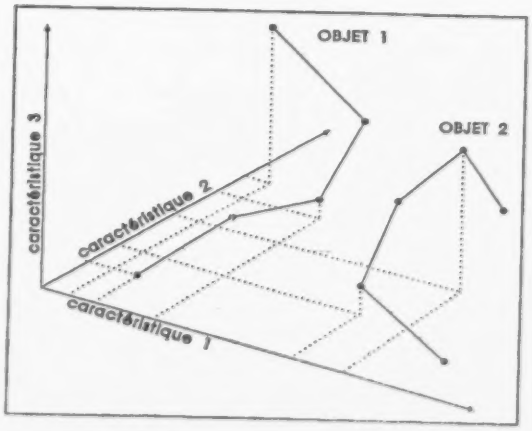


Figure 3: Trajectoires d'espace caractéristique pour des vecteurs à trois dimensions

objets. Ainsi, si on prend l'objet 1, on tient compte des images dans l'ordre suivant: 0, 15, 30, ... degrés. On suppose donc que le changement des vecteurs en fonction de la rotation hors plan subit par un objet entre angles d'apprentissage suit une relation linéaire. On peut visualiser les trajectoires d'espace caractéristiques en prenant des vecteurs de dimensions 3. La Figure 3 montre deux objets avec chacun cinq images apprises. Avec une rotation de 360 degrés, chaque objet serait représenté par une trajectoire constituée de segments de droites fermée dans l'espace caractéristique.

7.6. Reconnaissance

Une fois l'apprentissage terminé et les trajectoires dans l'espace caractéristiques bien déterminées, on peut faire la reconnaissance d'une image inconnue. On envoie cette image à travers le système pour en déterminer la position que sa signature occupe dans l'espace caractéristique. On mesure ensuite la plus petite distance entre cette inconnue et toutes les trajectoires d'espace caractéristique en mémoire. On attribut l'inconnue à l'objet possédant la trajectoire caractéristique la plus près. La Figure 4 illustre ceci dans un espace à deux dimensions pour permettre une meilleure visualisation. Dans ce cas, on voit que la distance de l'inconnu à l'objet 1 est plus petite que celle à l'objet 2. On conclue donc que l'inconnu est l'objet 1.

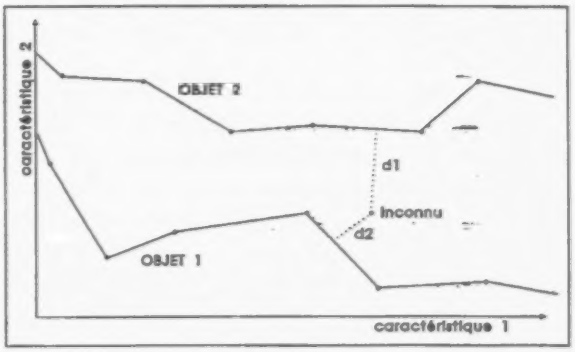


Figure 4: Reconnaissance par la distance minimale.

Pour calculer les distances minimales et tenir compte des droites reliant les images d'apprentissage d'un même objet, on utilise la multiplication scalaire que fait la projection de l'inconnue sur une droite connue. La Figure 5 montre le détail d'une telle opération. Une fois que la projection est trouvée, on calcule par trigonométrie la distance de l'inconnu à la droite.

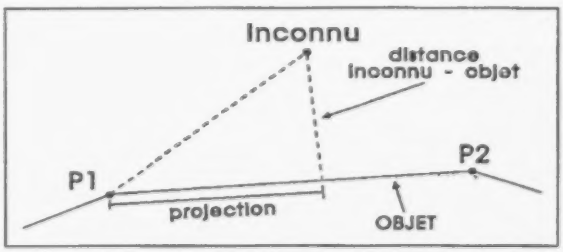


Figure 5: Calcul de la distance minimale à un objet

8. Résultats de la méthode FST NN

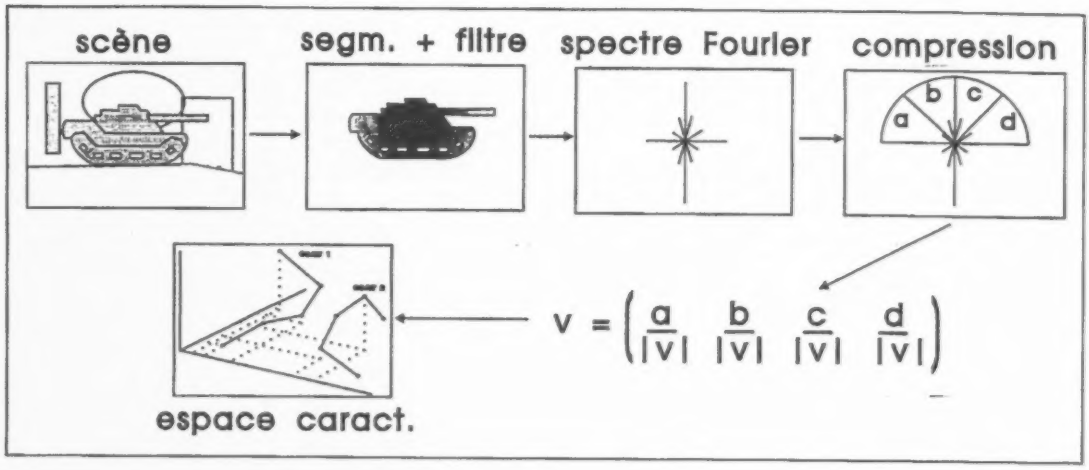


Figure 6: Schéma du montage utilisé lors du projet

Cette section présente les résultats de la simulation numérique de la méthode basée sur celle présentée par Neiberg et Casasent. Elle se divise en trois sections: l'invariance sous échelle, la robustesse face au bruit et la capacité de reconnaissance. Chacune des sections est un travail indépendant des deux autres. Elles permettent de voir la qualité du système face à un paramètre en particulier, soit un changement d'échelle, un ajout de bruit ou la reconnaissance d'images inconnues.

Avant d'étudier ces résultats, il est important de consulter le schéma de la méthode utilisée (Figure 6) et de le comparer avec celui de Neiberg et Casasent (Figure 2) car quelques petites modifications ont été apportées. Tout d'abord, on remarque que la binarisation des images n'a pas lieu. Nous avons préféré éliminer cette étape pour éviter la difficulté de rendre automatique et tout à fait objective l'application d'un seuil de binarisation. On remarque aussi l'ajout du filtrage des images à l'entrée du montage, l'ajout d'une normalisation des vecteurs de sortie résultant de la compression et l'élimination du réseau de neurones. Les détails de ces trois derniers changements seront donnés dans la section traitant de l'invariance sous échelle et celle de la capacité de reconnaissance.

Tout au long de ce chapitre, il y aura différentes formes de représentations des résultats qu'il est utile d'expliquer avant de les utiliser. Tout d'abord il faut comprendre ce qu'on appelle la reconnaissance et la classification. La reconnaissance fait appel à la capacité d'associer une image inconnu au bon objet. La classification fait plutôt appel à la capacité d'associer une image inconnu à la bonne classe d'objets (char d'assaut ou transporteur de troupes). C'est ce dernier critère qu'il est important d'optimiser dans le présent travail.

Une autre représentation des résultats de reconnaissance est le graphique de la distance vs le point de vue (exemple Figure 9). Ce type de graphique présente en ordonné la distance minimale de chacune des trajectoires caractéristiques à une image inconnue. En abscisse, nous avons l'image inconnue qui est la banque complète d'images d'un objet (le chief pour la Figure 9). Ainsi les chiffres en abscisse

représentent le point de vue en rotation hors plan de l'objet entré comme inconnu. Si on regarde sur le graphique de la figure 8 à la position 35 en abscisse (35 degrés de rotation hors plan), cela nous indique en ordonnée la distance du chief 35 aux cinq trajectoires d'espace caractéristique en mémoire. La distance la plus près de zéro fixe l'objet attribué à l'inconnu. Idéalement nous aurions donc une valeur de zéro pour tous les points de vue du chief et une valeur très grande pour les autres objets puisque les inconnus utilisés appartiennent au chief.

Finalement, il y a la présentation sous forme d'histogramme de plage d'invariance d'échelle. La Figure 24 en est un exemple. On retrouve en abscisse une image à étudier et en ordonnée la plage d'invariance sous échelle de cette image. La plage d'invariance indique le plus grand écart de changement d'échelle qu'il est possible de faire subir à une image pour qu'elle demeure reconnue par le système. Par exemple la plage de 5.2 indique que l'objet peut être reconnu dans l'intervalle de 1.3 à 0.25 fois sa grosseur initiale.

8.1. Invariance sous échelle

Il est important d'étudier l'invariance sous échelle de l'algorithme pour s'assurer que la discrimination se fait bien sur la différence entre les objets et non sur leur différence d'échelle. Dans notre banque d'images il y a des objets de dimensions différentes. La différence la plus grande est probablement entre le chief et le btr152. Cette différence est environ d'un facteur 2. Il faudrait donc idéalement avoir au minimum une plage d'invariance de deux. La variation d'échelle des images d'entrée peut aussi être causée par un éloignement ou un rapprochement entre l'objet et la caméra. Comme la banque d'images reçue ne contient pas d'objets à des échelles différentes, il faut procéder à un changement d'échelle numérique pour simuler cet effet. Il faut faire attention que ce changement d'échelle n'occasionne pas d'artéfacts dans les images. La méthode de changement d'échelle utilisée sera l'interpolation linéaire. Cette méthode est probablement différente de ce qui se produit en réalité, mais elle donne tout de même une bonne approximation, surtout pour la réduction de dimension puisqu'on a pas à extrapoler d'information. De plus, si les images utilisées respectent les critères d'échantillonnage, la diminution de l'objet d'un facteur 2 ou 4 est exacte peu importe la méthode de réduction de dimension utilisée. Ainsi, l'image dont la taille est réduite numériquement par un tel facteur contient la même information que celle captée à cette échelle. La réduction maximale des images est de 0.25 et le grandissement maximal est de 1.3. Cette valeur est la limite du grossissement du plus gros objet sans que ce dernier ne déborde de la matrice qui le contient. La plage d'invariance étudiée est ainsi d'un facteur 5.

8.1.1 Normalisation des vecteurs

Lors de l'utilisation de la compression en secteur, il est important de faire la normalisation des vecteurs de sortie pour vraiment avoir l'invariance sous échelle. Le changement d'échelle ne fait que dilater ou compresser le spectre de Fourier radialement ce qui conserve la proportion dans les sections mais change l'énergie totale. En normalisant on élimine l'information de l'énergie totale pour ne tenir compte que de la proportion.

Nous montrerons ici les résultats pour des vecteurs non normalisés et pour des vecteurs normalisés de façon à mesurer l'effet de ce paramètre. Les graphiques des Figure 9 à Figure 13 montrent la reconnaissance globale du système avec des vecteurs non-normalisés. Les résultats de la reconnaissance et de la classification sont présentés au tableau 1.

VÉHICULES	CHIEF	LÉOPARD	BTR152	GRIZZLY	PT76	TOTAL
mauvaise reconnaissance	2	1	3	1	4	11
mauvaise classification	0	1	0	0	1	2

Tableau 1: Analyse de la reconnaissance et de la classification pour des vecteurs non-normalisés.

Les résultats de la classification sont très bons, soit de 99.4% (2 sur 360 images) si on tient compte de toute la banque d'images et de 99.2% (2 sur 240 images) si l'on ne tient pas compte des 120 images de l'apprentissage. Si on regarde maintenant l'invariance sous échelle, on remarque que la situation n'est pas aussi intéressante que la précédente. Par les graphiques des figures 23 à 27 on peut voir que dans les images étudiées il y en a 1/3 qui ne peuvent pas être changées ou très peu changées (plage environ de 1) sans éviter des erreurs de classification. De plus, 18% seulement peuvent être reconnues si elles sont de plus de la moitié (plage de 2X) de leur grosseur initiale. Ce résultat doit absolument être amélioré pour respecter les écarts d'échelle de notre banque d'images.

Lorsqu'on normalise les vecteurs de sortie de la compression, les résultats de la reconnaissance ne changent pas beaucoup mais l'invariance sous échelle s'améliore grandement. Les résultats de reconnaissance et de classification sont illustrés au graphique des Figure 14 à Figure 18 et sont résumés au tableau 2.

VÉHICULES	CHIEF	LÉOPARD	BTR152	GRIZZLY	PT76	TOTAL
mauvaise reconnaissance	1	2	3	4	3	13
mauvaise classification	1	2	0	0	0	3

Tableau 2: Analyse de la reconnaissance et de la classification pour des vecteurs normalisés

Les résultats de la classification sont encore très bons, soit de 99.2% si on tient compte de toute la banque d'images et de 98.75% si l'on ne tient pas compte des images de l'apprentissage. Si on étudie maintenant l'invariance sous échelle, on remarque que la situation s'est améliorée. Par les graphiques des figures 23 à 27 on peut voir que pour les images étudiées il y a maintenant 10% des images qui ne peuvent pas être changées ou très peu changées pour éviter qu'elles ne puissent être reconnues et environ 71% peuvent être reconnues si elles sont modifiées de plus de la moitié de leur grosseur initiale. La normalisation augmente donc de beaucoup l'invariance sous échelle sans trop détériorer la classification des objets. Il demeure encore un pourcentage trop élevé d'images qui ne sont pas reconnues s'il y a de petits changements d'échelle.

8.1.2 Filtrage moyen

Lorsque la taille de l'objet est réduite, la taille de sa transformée de Fourier augmente et risque de dépasser le cadre de la matrice dans les hautes fréquences. Si cela ce produit, nous avons un phénomène de repliement spectral (« aliasing ») qui déforme l'information des hautes fréquences. Une fois que ceci se produit, il n'y a aucune façon de retrouver exactement l'information initiale. Cet aliasing est d'autant plus important que la réduction d'échelle est grande. Il en résulte une erreur dans le traitement ultérieur des images. Il faut donc filtrer les images pour s'assurer qu'elle sont correctes avant le changement d'échelle.

Pour filtrer les images, nous utilisons un filtre moyen. Il faut filtrer les images avant de faire un changement d'échelle et une transformée de Fourier. Ceci permet donc de s'assurer que le spectre de puissance ne débordera pas sa matrice lors du changement d'échelle.

L'application de ce filtre s'effectue par la convolution de l'image avec une matrice de moyennage. Dans notre cas, on effectue deux fois la convolution avec une matrice 3X3 qui prend la moyenne des neufs pixels la composant et remplace la valeur centrale où tombe la matrice. Le développement mathématique de ce filtre est le suivant:

$$g(x) = f(x) * m(x) = TF\{TF\{f(x)\} \cdot TF\{m(x)\}\} = TF\{F(u) \cdot M(u)\}$$

où g(x) est l'image résultante, f(x) est l'image à filtrer, m(x) est le filtre. Comme on le sait, la transformée de Fourier (TF) d'une convolution de fonctions est la TF de la multiplication de la TF de ces fonctions (théorème de convolution). Si on regarde l'allure de la TF de m(x) (Figure 7) on s'aperçoit que c'est un filtre passe-bas. Ainsi les hautes fréquences causant le repliement spectral sont coupées.

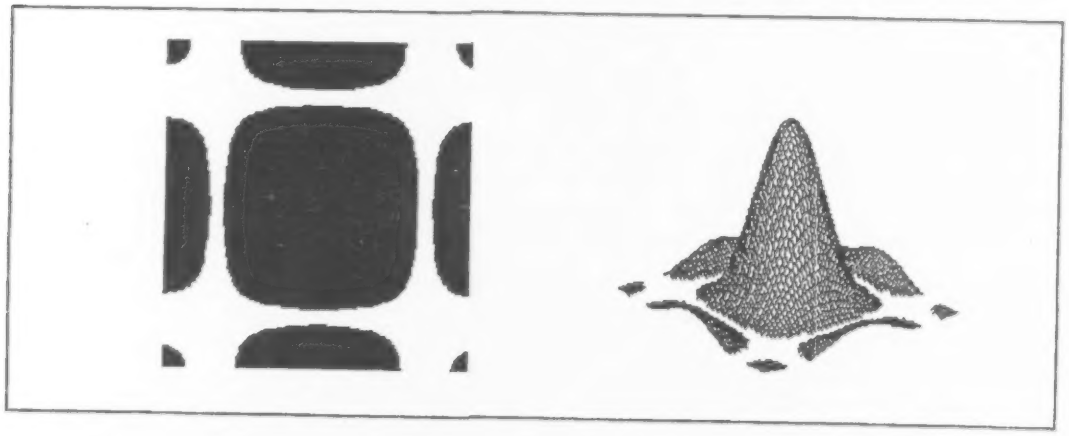


Figure 7: Allure du filtre moyen dans le plan de Fourier. A gauche vue de dessus et à droite vue en 3-D

Les résultats de reconnaissance du système avec filtrage moyen sont présentés dans les Figure 19 à Figure 23 et sont résumés dans le tableau 3.

VÉHICULES	CHIEF	LÉOPARD	BTR152	GRIZZLY	PT76	TOTAL
mauvaise reconnaissance	2	5	3	5	3	18
mauvaise classification	2	5	0	0	1	8

Tableau 3: Analyse de la reconnaissance et de la classification pour des vecteurs normalisés et filtrés

Les résultats de la classification sont un peu moins bons que les précédents, soit de 97.8% si on tient compte de toute la banque d'images et de 96.7% si l'on ne tient pas compte des images de l'apprentissage. Si on étudie maintenant l'invariance sous échelle, on remarque que la situation s'est encore améliorée. Par les graphiques des Figure 24 à Figure 28 on peut voir que pour les images étudiées il n'y a plus qu'environ 2% des images qui ne peuvent pas être changées ou très peu changées pour éviter qu'elles ne puissent être reconnues et environ 94% peuvent être reconnues si elles sont modifiées de plus de la moitié de leur grosseur initiale. On voit aussi que 70% sont reconnues malgré un changement d'échelle d'un facteur 4, comparativement à 59% pour les images non filtrées et 0% pour les images non filtrées et non normalisées.

8.1.3 La qualité de l'invariance sous échelle

On peut se demander pourquo l'invariance sous échelle n'est pas parfaite malgré la normalisation et le filtrage. Les causes sont de plusieurs ordres. Il y a tout d'abord l'erreur occasionnée par l'utilisation de pixels carrés pour faire la simulation du compresseur en secteurs. Toute la région centrale des spectres de Fourier utilisés souffre d'un manque de résolution à cause de la géométrie cartésienne des détecteurs par rapport à la géométrie circulaire du spectre de Fourier (voir Figure 8). Cette région étant celle contenant beaucoup d'énergie, sa répercussion est grande dans le cas d'une erreur. Il y a aussi le filtre moyen 3X3 utilisé qui n'est pas un filtre parfait, c'est à dire qui ne coupe pas complètement les hautes fréquences dans les directions horizontale et verticale du spectre de Fourier. Or, à cause de la géométrie des objets utilisés, ce sont justement dans ces régions que sont situées les plus hautes fréquences. Si le filtre ne coupe pas bien ces directions, il y a de fortes chances d'avoir encore du repliement spectral lors du changement d'échelle malgré le filtrage.

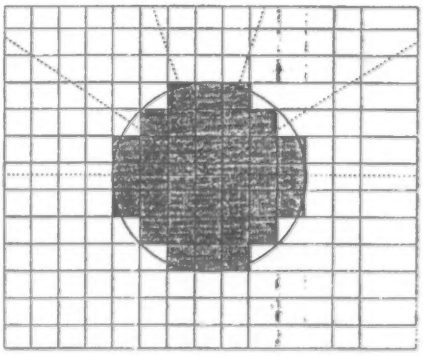


Figure 8: Manque de résolution au centre du spectre à cause des pixels carrés

En étudiant les résultats précédents, on peut aussi se demander pourquoi les plages d'invariances varient autant d'une image à une autre. Ceci peut être expliqué par les problèmes de résolution, tels

qu'expliqués au paragraphe précédent, combinés à la qualité de la reconnaissance de l'image inconnue par le système. Par exemple, pour le léopard, l'invariance sous échelle est la moins bonne pour trois points de vue: 150, 180 et 310 degrés. Si on regarde la reconnaissance de ces images par rapport aux sept autres images étudiées pour l'invariance sous échelle (Figure 20), on remarque qu'elles sont celles qui présentent la moins bonne discrimination. Le même phénomène s'observe pour le btr152 à 35 degrés et 235 degrés (Figure 21). Si on prend le temps d'analyser la corrélation entre la plage d'invariance et la qualité de la discrimination, on s'aperçoit que plus la discrimination est forte lors de la reconnaissance, plus l'invariance sous échelle couvre une grande plage. Ce phénomène est causé par les erreurs de simulations décrite au paragraphe précédent.

On peut faire un résumé de ce qui vient d'être expliqué dans cette section à l'aide du tableau 4. On peut ainsi facilement voir la faible décroissance de la qualité de la classification par rapport à la nette amélioration de l'invariance sous échelle selon les opérations apportées dans le système. Nous utiliserons le filtrage moyen et la normalisation dans les deux prochaines sections pour s'assurer d'avoir la meilleure invariance sous échelle possible.

	taux classification		Invariance sous échelle		
	apprentissage + test	test seulement	sans changements	de 1X à 2X	2X et +
sans filtre + sans normalisation	99.4%	99.2%	33%	49%	18%
sans filtre + avec normalisation	99.2%	98.8%	10%	18%	72%
avec filtre + avec normalisation	97.8%	96.7%	2%	4%	94%

Tableau 4: Résumé des analyses de classification vues dans le chapitre.

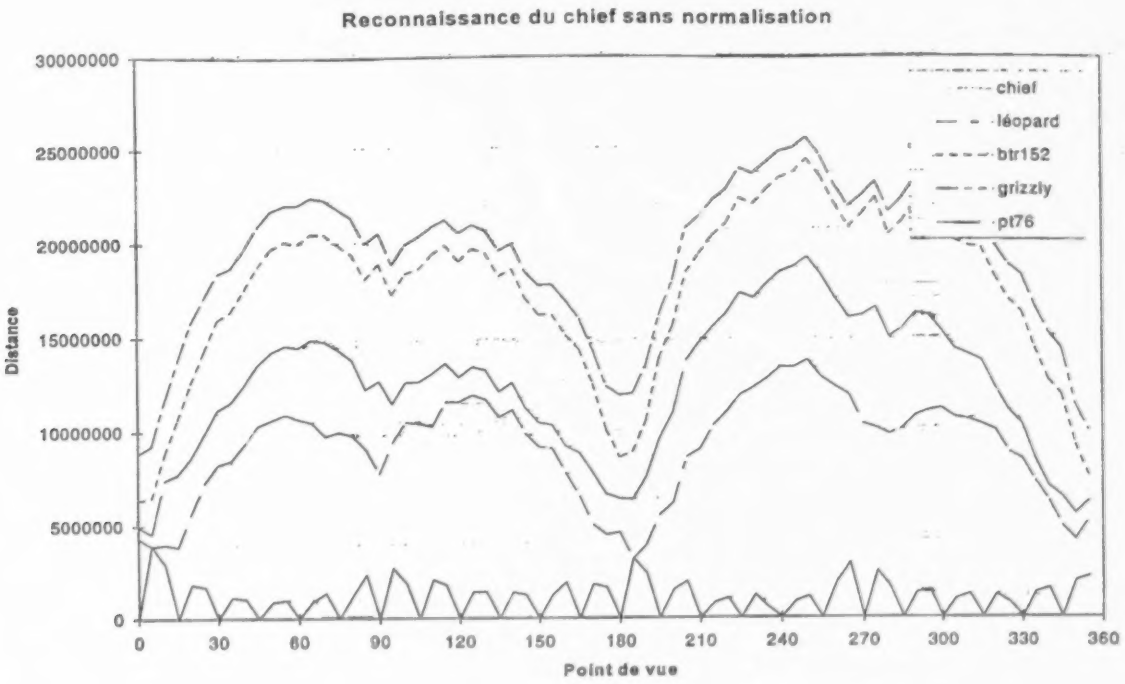


Figure 9

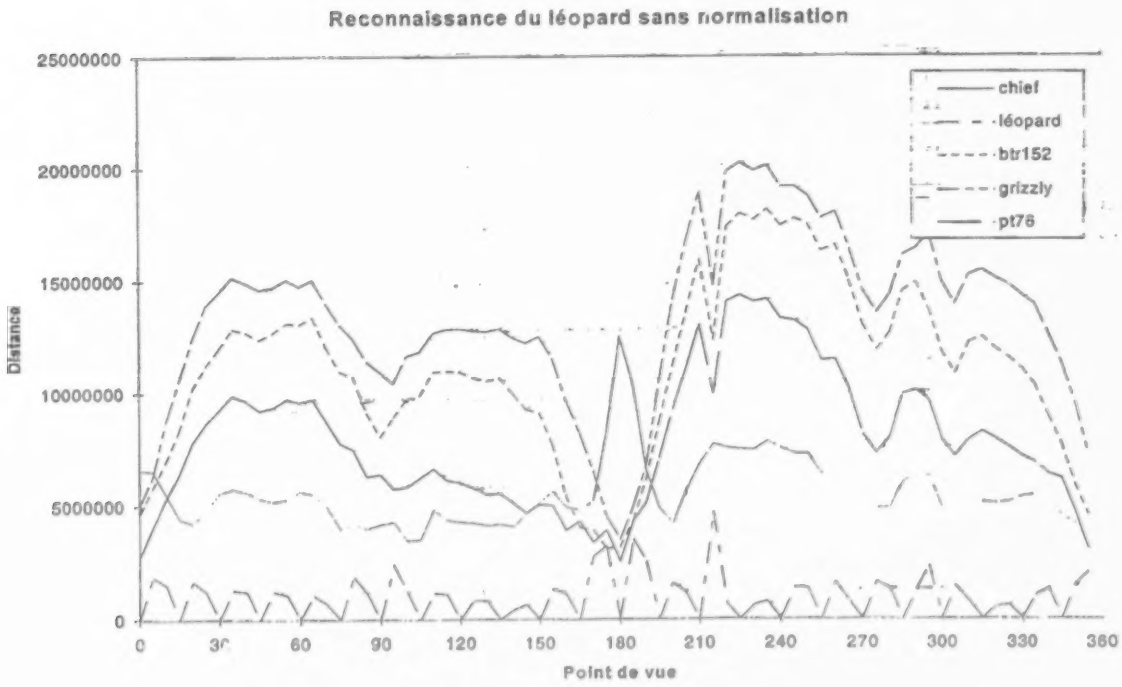


Figure 10

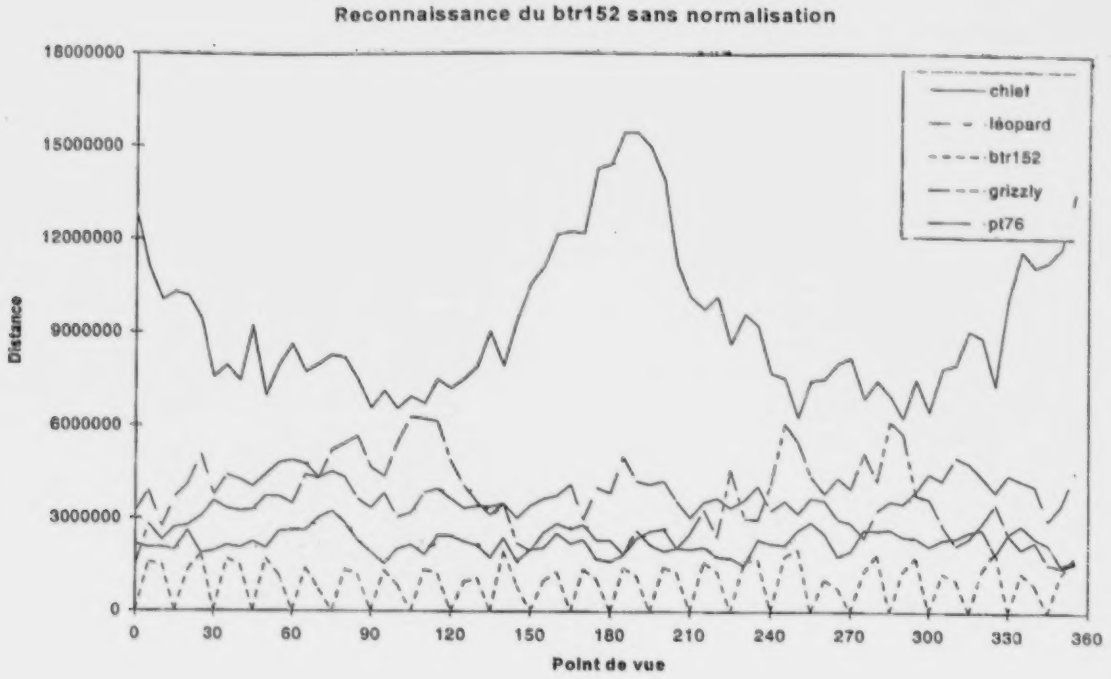


Figure 11

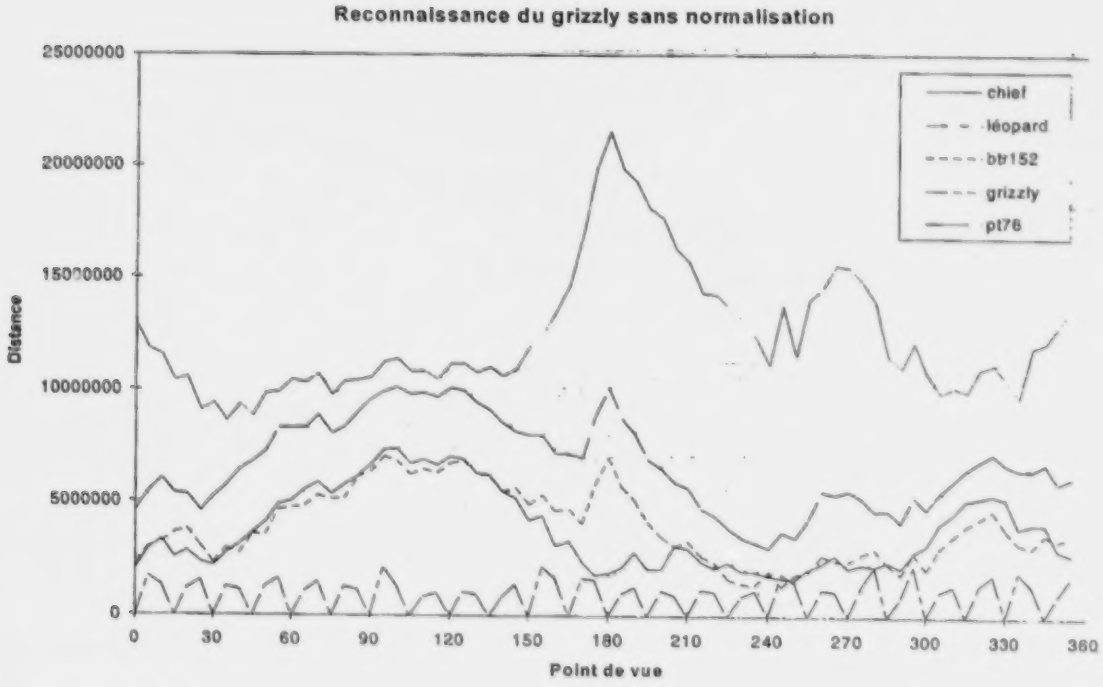


Figure 12

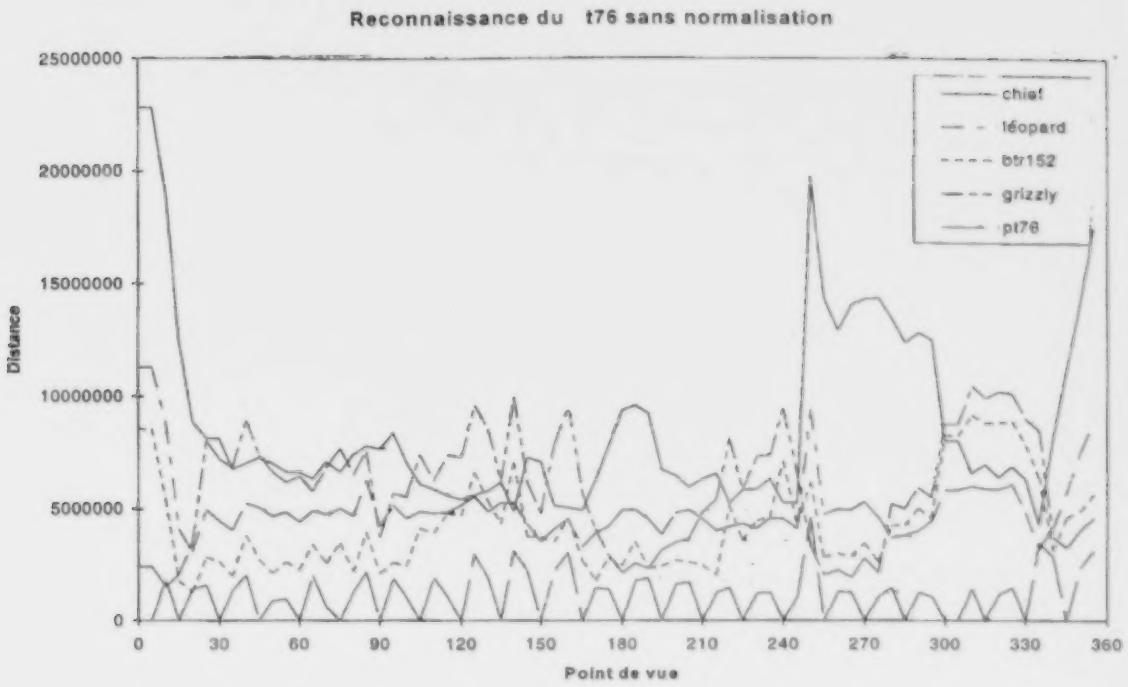

Reconnaissance du t76 sans normalisation
25000000
20000000
15000000
10000000
5000000
0
Distance
0
30
60
90
120
150
180
210
240
270
300
330
360
Point de vue
chief
léopard
btr152
grizzly
pt76

Figure 13

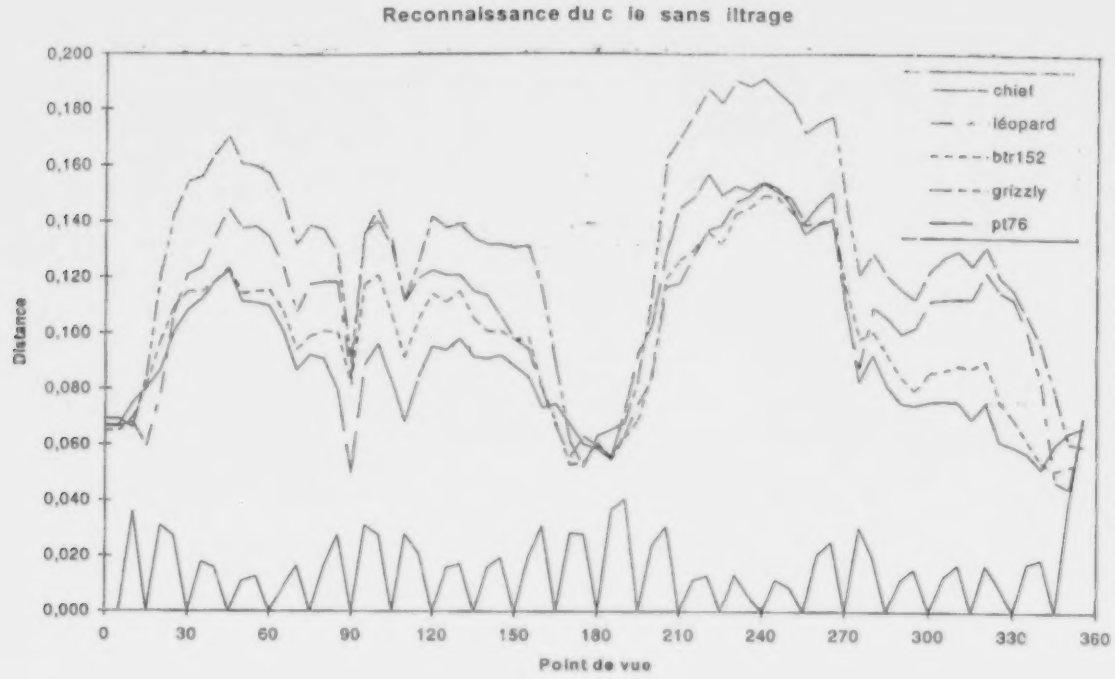


Figure 14

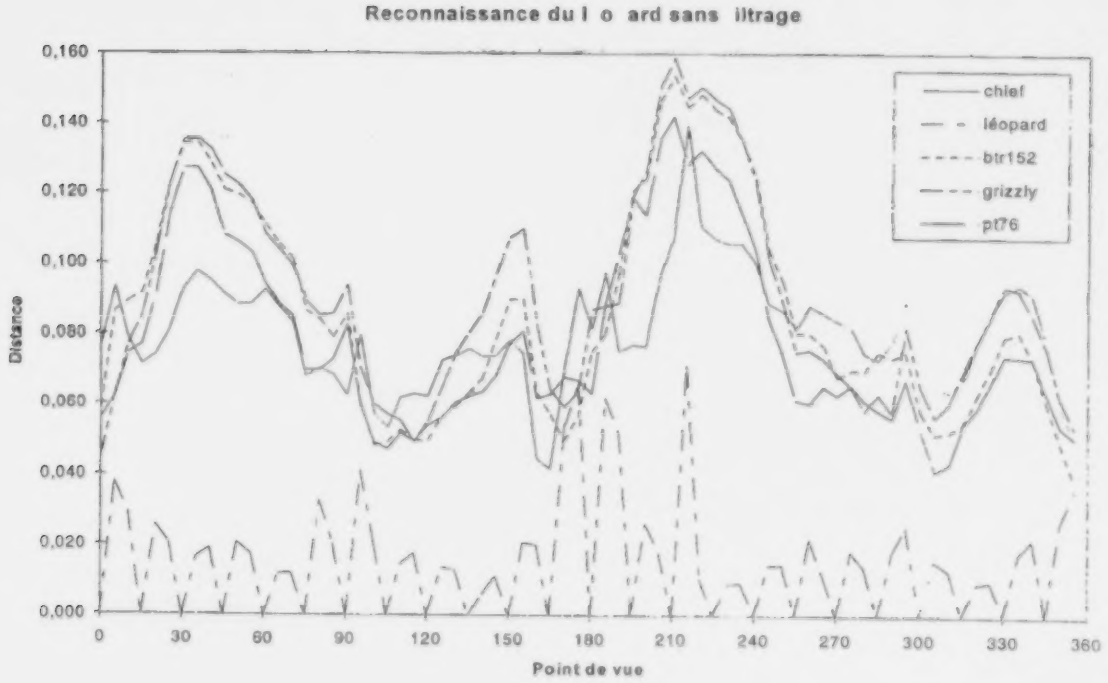


Figure 15

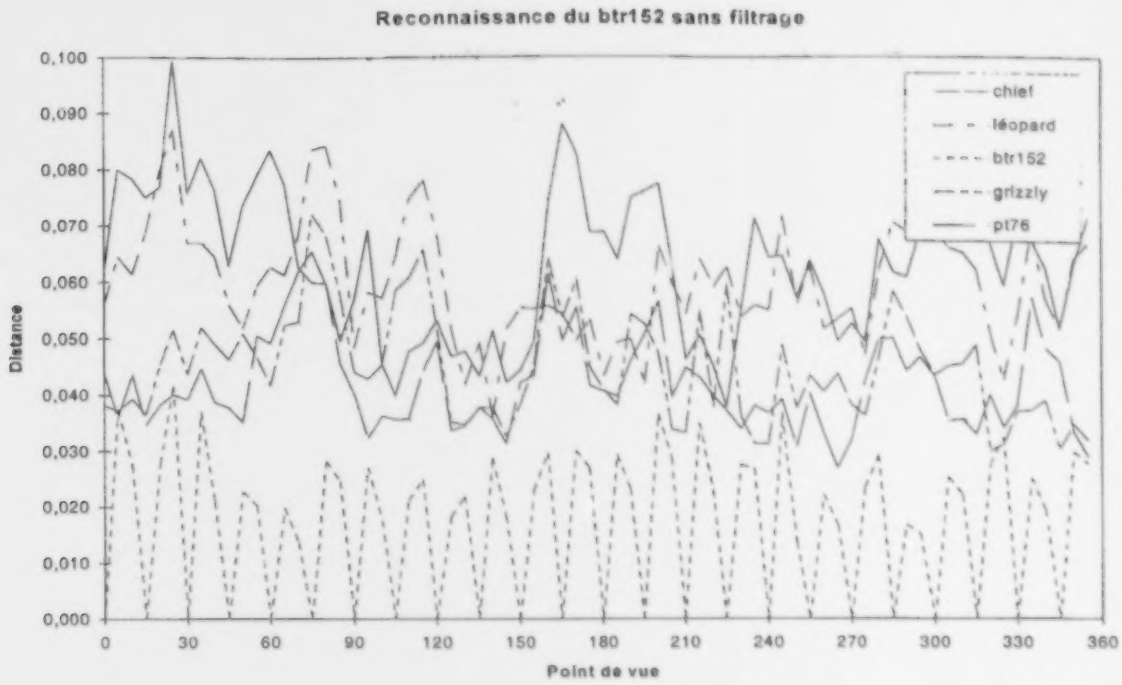


Figure 16

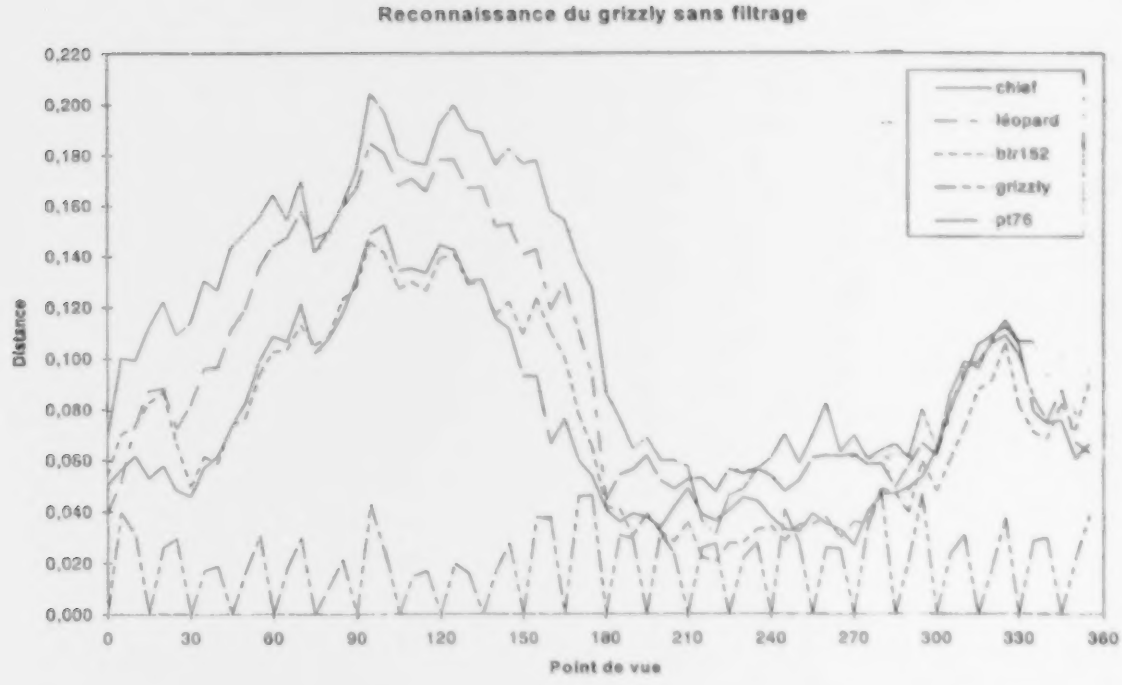


Figure 17

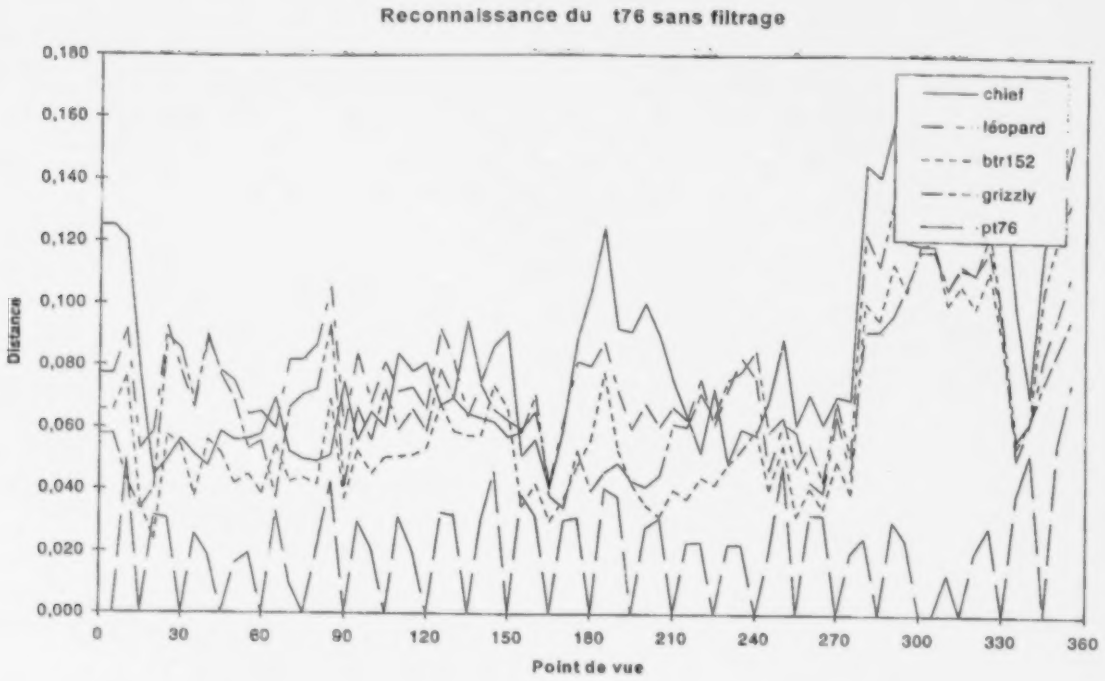

Reconnaissance du t76 sans filtrage
Distance
0,180
0,160
0,140
0,120
0,100
0,080
0,060
0,040
0,020
0,000
0
30
60
90
120
150
180
210
240
270
300
330
360
Point de vue
chief
léopard
btr152
grizzly
pt76

Figure 18

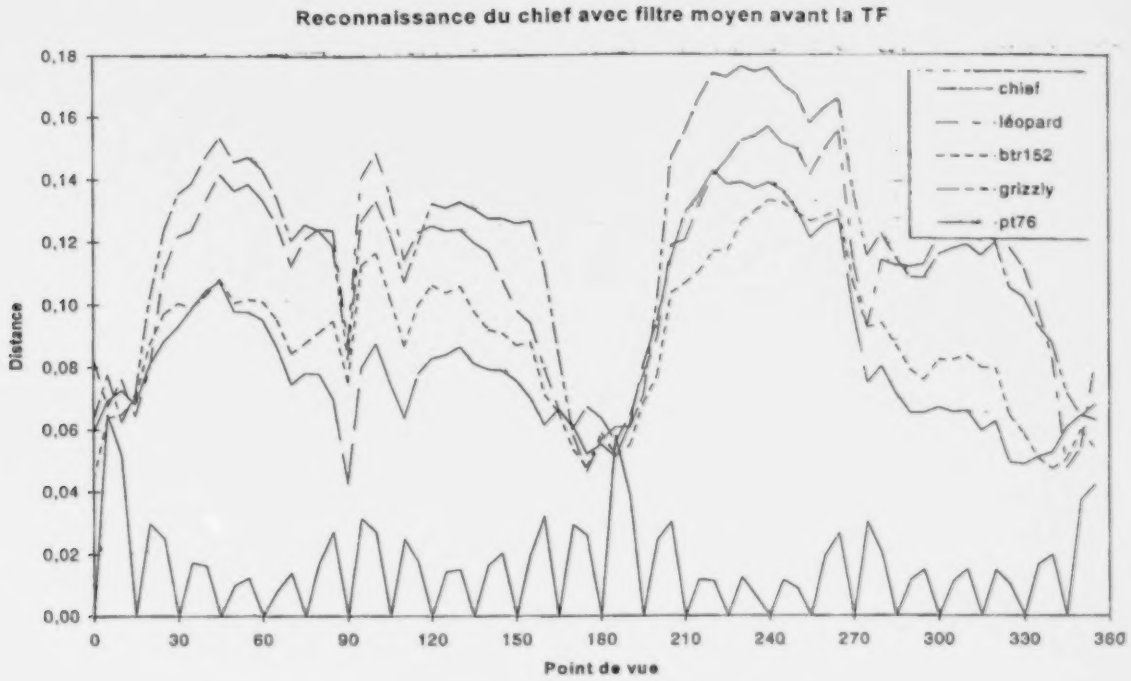


Figure 19

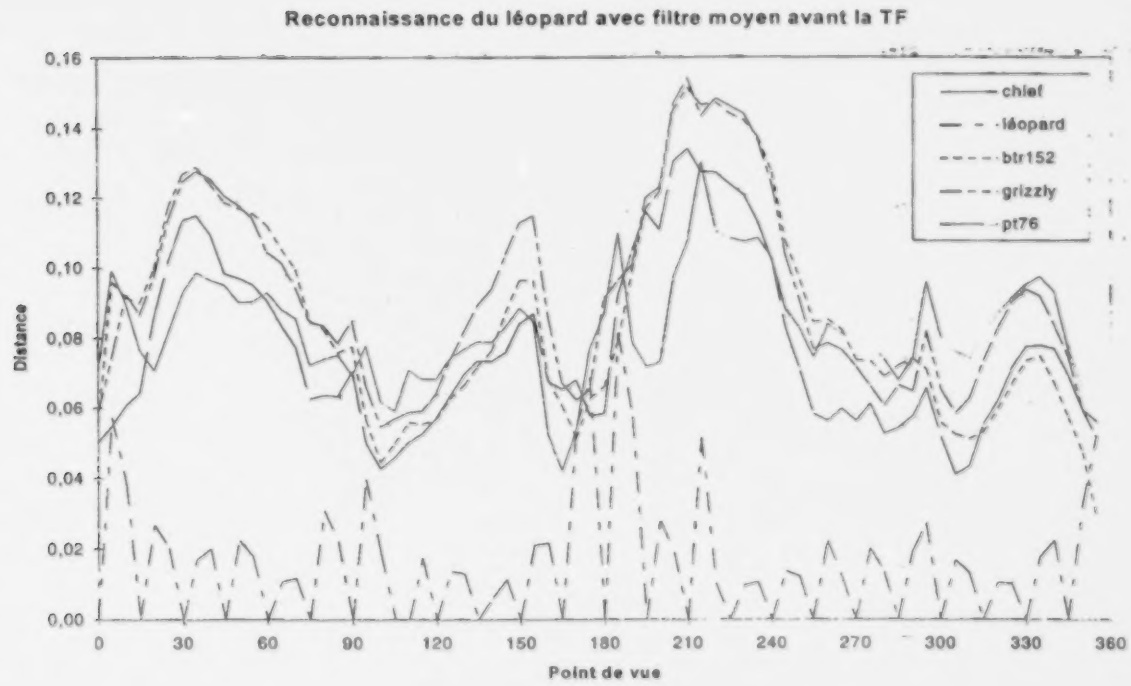


Figure 20

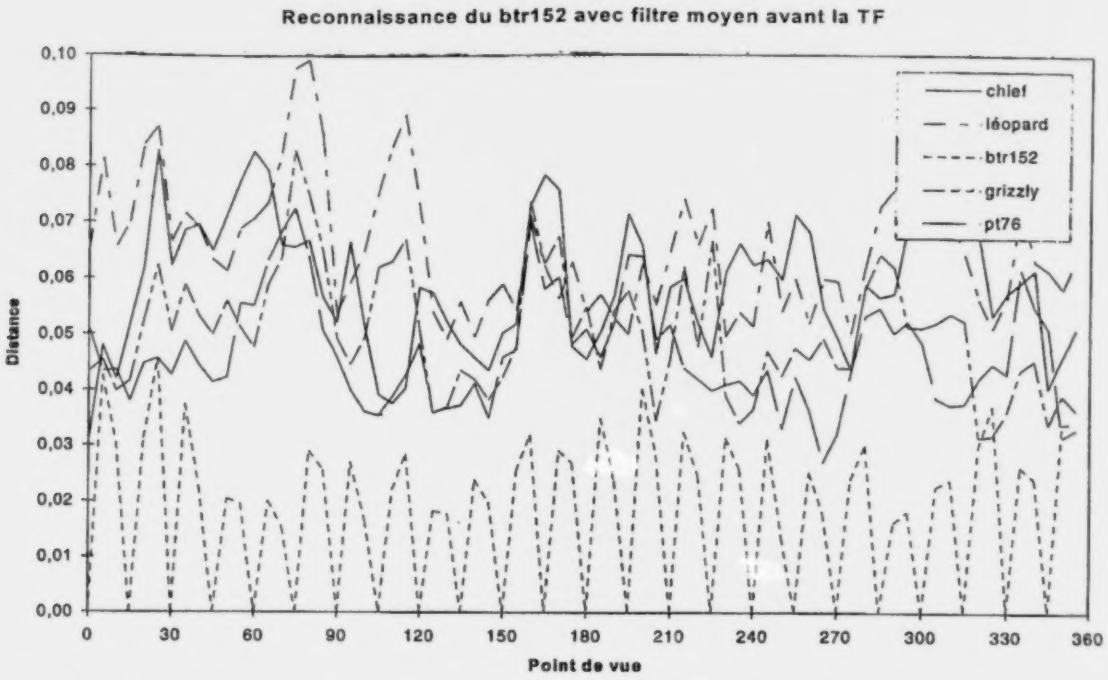


Figure 21

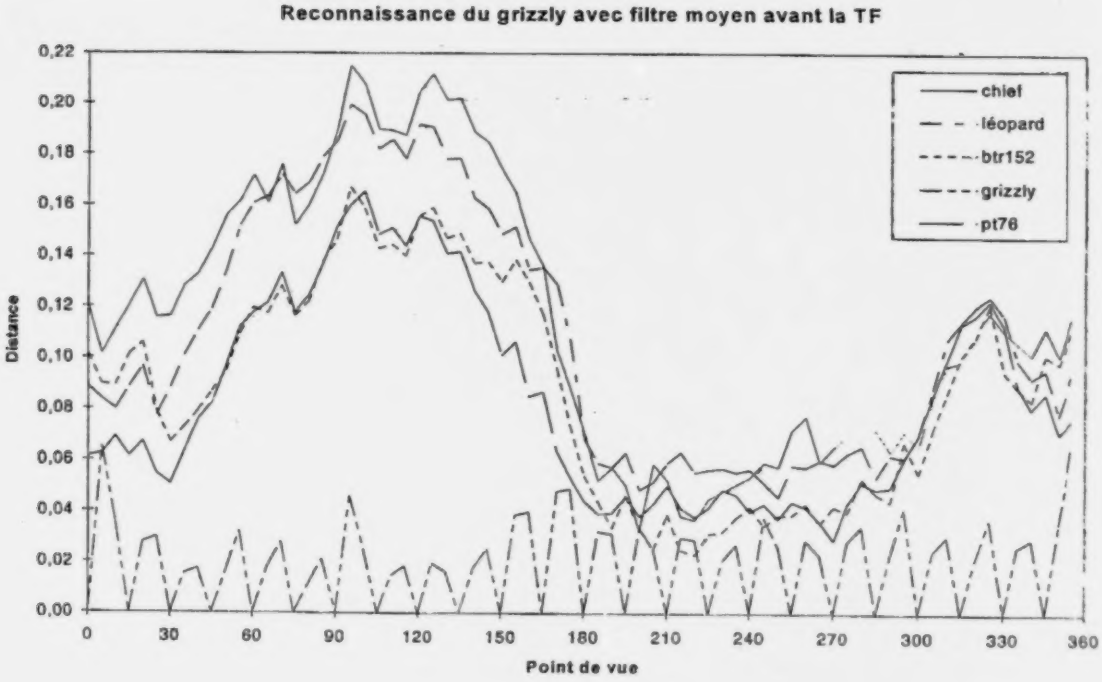


Figure 22

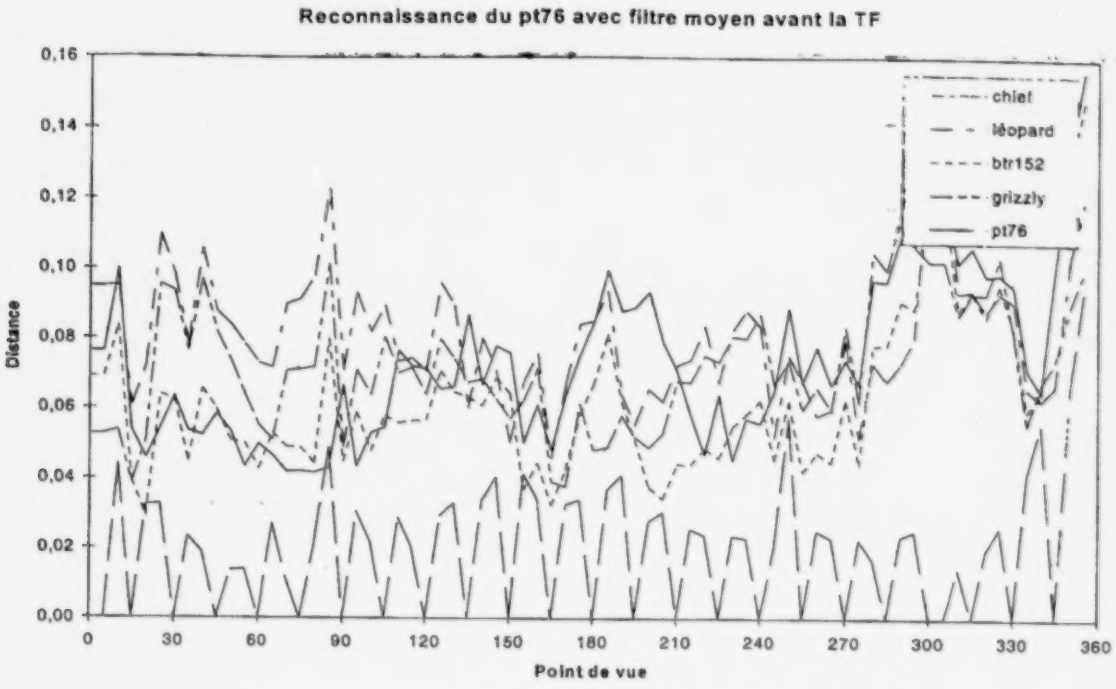

Reconnaissance du pt76 avec filtre moyen avant la TF
0,16
0,14
0,12
0,10
0,08
0,06
0,04
0,02
0,00
Distance
0
30
60
90
120
150
180
210
240
270
300
330
360
Point de vue
chief
léopard
btr152
grizzly
pt76

Figure 23

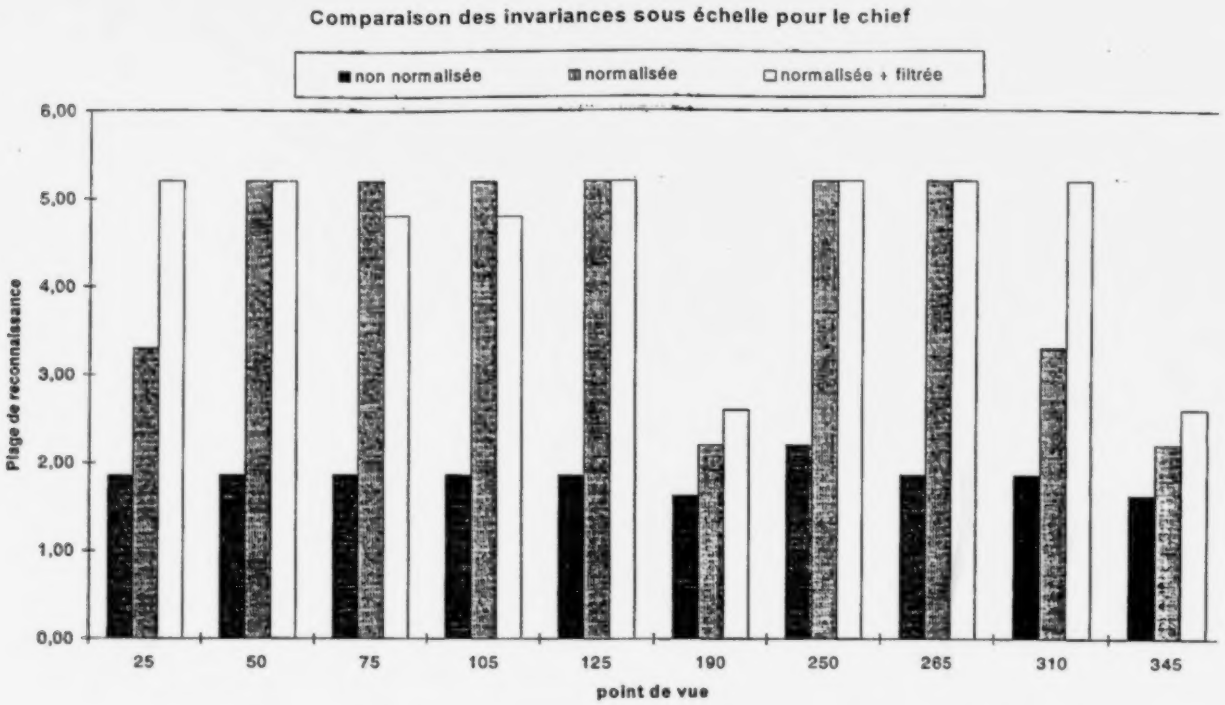


Figure 24

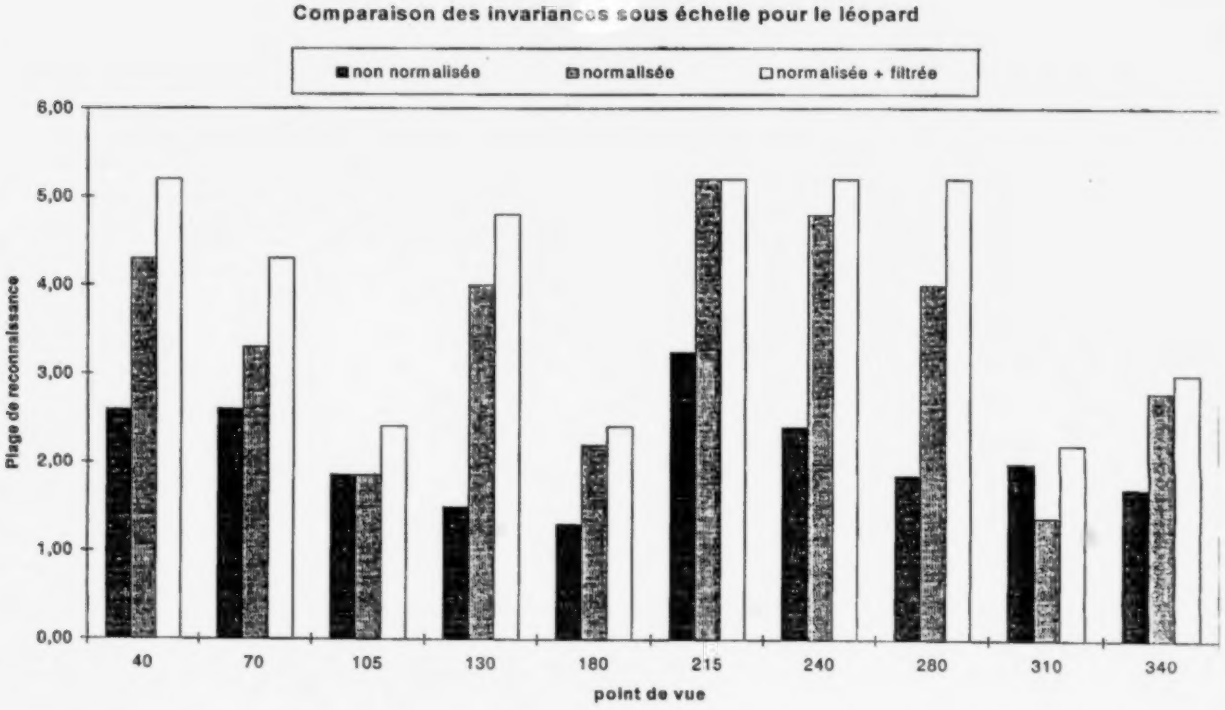


Figure 25

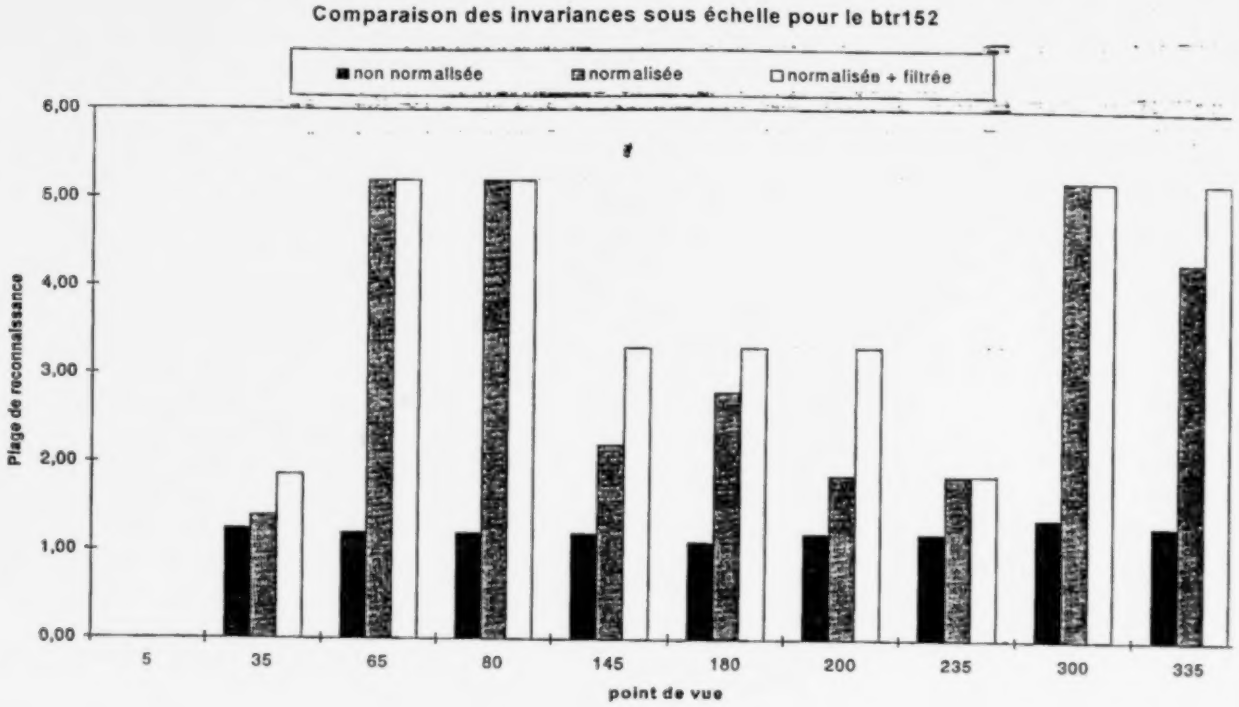


Figure 26

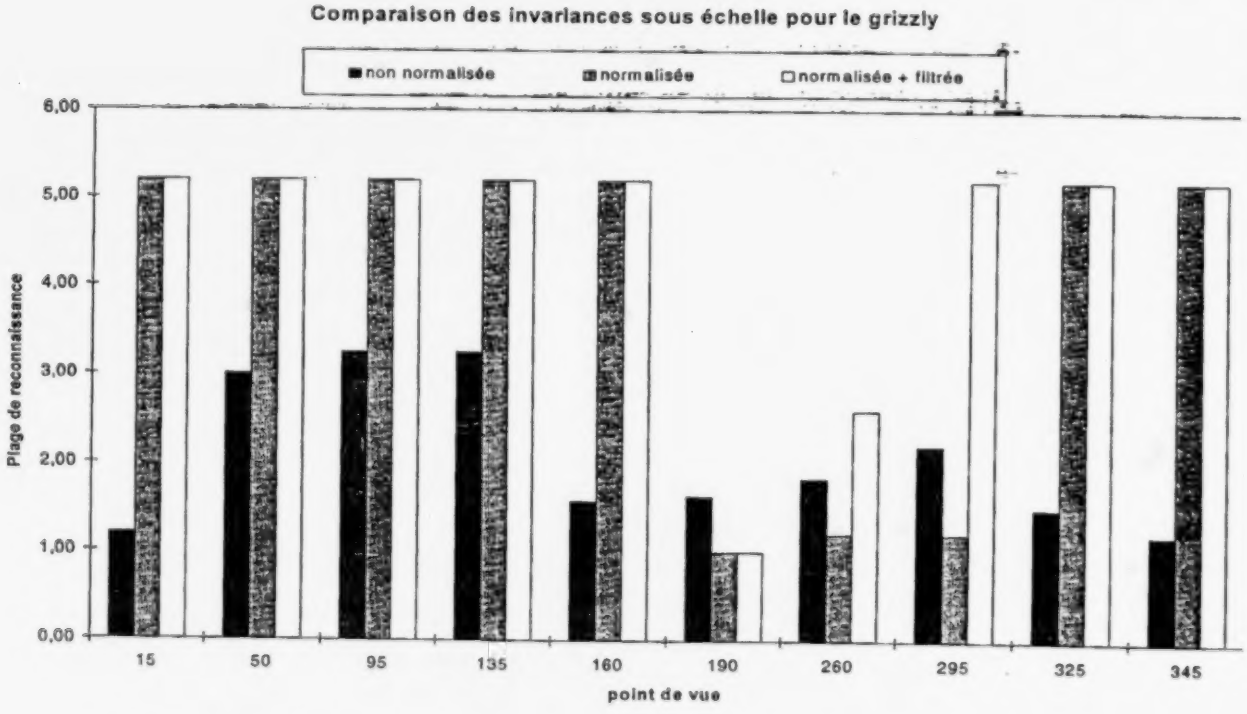


Figure 27

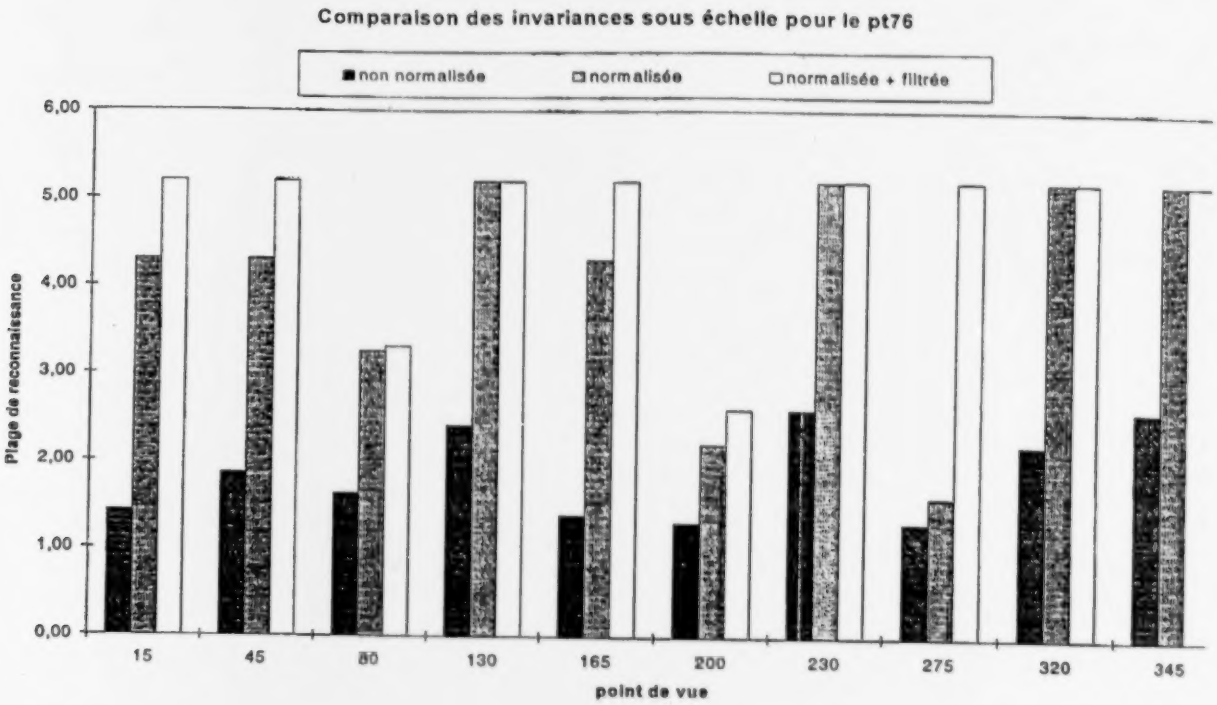

Comparaison des invariances sous échelle pour le pt76
non normalisée
normalisée
normalisée + filtrée
Plage de reconnaissance
6,00
5,00
4,00
3,00
2,00
1,00
0,00
15
45
80
130
165
200
230
275
320
345
point de vue

Figure 28

8.2. Robustesse face au bruit

Il est intéressant d'évaluer quelle est la robustesse de l'algorithme face au bruit. Nous avons étudié l'effet du bruit qui se présente majoritairement sous deux formes: le bruit blanc et le bruit de fond. Le bruit blanc est l'approximation plus ou moins exacte du bruit occasionné par les fluctuations thermiques de l'image lors de l'acquisition avec la caméra infrarouge. Ce type de bruit est comme la neige à l'écran d'un téléviseur lorsque que le poste n'est pas bien reçu. La caractéristique du bruit blanc est son spectre uniforme, d'où vient son nom. L'amplitude à toutes les fréquences est sensiblement la même et sa phase est aléatoire. Ce bruit est ajouté uniformément sur toute la région de l'image à capter. Cela se traduit par l'addition du spectre du bruit au spectre de l'image. Le deuxième type de bruit est le bruit de fond qui correspond à un décor par dessus lequel notre objet à reconnaître se retrouve. C'est le décor dans lequel on retrouve l'objet étudié. Ici aussi le spectre total sera le spectre de l'objet additionné au spectre du fond.

Comme nous utilisons le spectre de puissance, le résultat n'est pas simplement l'addition du spectre de puissance du bruit au spectre de puissance de l'image. Ceci est montré dans l'équation suivante:

$$\left|TF\{O+B\}\right|^2 = \left|TF\{O\}\right|^2 + \left|TF\{B\}\right|^2 + 2\cdot\left|TF\{O\}\right|\cdot\left|TF\{B\}\right|\cdot\cos(\Delta\varphi)$$

où O et B sont les fonctions de l'objet et du bruit en fonction de x et y, $\Delta\varphi$ est la différence de phase entre la TF de l'image et la TF du bruit en fonction de x et y. On trouve donc que le spectre de puissance de l'image avec du bruit est le spectre de puissance de l'image additionné du spectre de puissance du bruit auquel s'ajoute un terme d'interférence dépendant de la différence de phase entre le bruit et l'image.

On remarque par cette équation que lorsque le bruit est d'énergie très faible, son spectre de puissance s'approche de zéro et est très faible par rapport à celui de l'image. Ainsi, le spectre de puissance total ressemble au spectre de puissance de l'objet sans bruit. Si le bruit est très élevé par rapport à l'image, c'est le contraire qui se produit. Le spectre de puissance total tend à ressembler à celui du bruit seul. On voit donc que plus le bruit sera élevé, plus la signature de la compression sera déformée par rapport à signature de l'objet seul. Mais il n'y a pas que l'intensité du bruit qui importe, il y a aussi la façon dont est répartie cette énergie dans le bruit. En effet, si le bruit a la même sortie du compresseur que l'image à reconnaître (c'est-à-dire la même répartition d'énergie dans les secteurs du compresseur), l'ajout du bruit à cet image aura un faible effet. Au contraire, moins les distributions d'énergie se ressemblent, moins l'énergie du bruit a besoin d'être élevée pour dégrader la reconnaissance.

Pour cette étude, nous utiliserons trois images de la banque fournies pour le projet. Les objets choisis seront le léopard vu à 90 degrés de rotation hors plan, le btr152 à 180 degrés et le pt76 à 270 degrés. Les signatures sont regroupées dans le tableau 5.

Vecteurs de sortie du compresseur en secteur pour les objets								
Objets	**Vecteurs**							
léopard	0.267	0.203	0.185	0.161	0.185	0.184	0.250	0.510
(90 degrés)	0.465	0.218	0.176	0.162	0.130	0.137	0.157	0.237
btr152	0.314	0.184	0.151	0.142	0.156	0.184	0.235	0.457
(180 degrés)	0.466	0.239	0.185	0.163	0.149	0.145	0.170	0.307
pt76	0.256	0.152	0.134	0.143	0.158	0.183	0.242	0.463
(270 degrés)	0.548	0.223	0.190	0.177	0.147	0.145	0.147	0.246

Tableau 5

8.2.1 Bruit Blanc

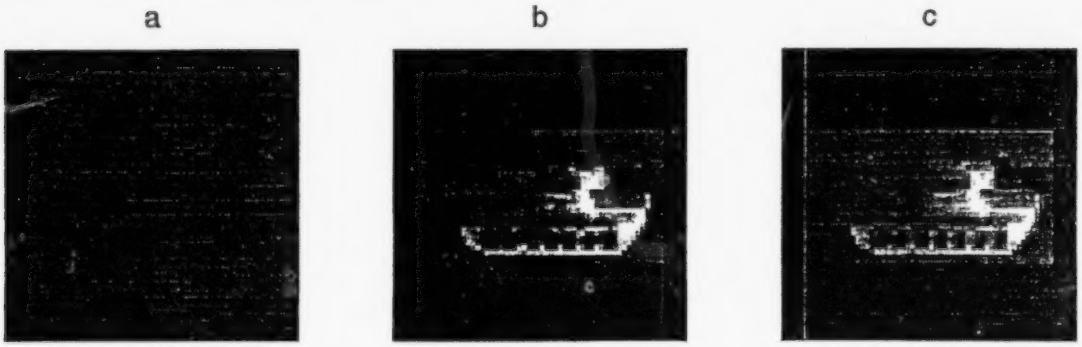


Figure 29: a) Bruit blanc seul (normalisé), b) Objet seul, c) Objet avec du bruit de moyenne 100 et d'écart type de 3.

Les graphiques des Figure 31 à Figure 33 montrent les résultats de l'étude sur le bruit blanc. On utilise trois objets, soit le léopard vu à 90 degrés, le btr152 à 180 degrés et le pt76 à 270 degrés auxquels on ajoute du bruit blanc. Le bruit blanc utilisé est réparti aléatoirement dans la matrice d'observation. Les paramètres qui caractérisent ce type de bruit sont la moyenne et l'écart type. La moyenne détermine la valeur centrale de la courbe normale, donc indique l'énergie contenu dans le bruit, tandis que l'écart type contrôle l'étalement de la répartition normale. Dans notre étude, nous utilisons un écart type fixé à 3 et une moyenne variant entre 10 et 110. On étudie l'effet du bruit sur la reconnaissance des objets par la distance de reconnaissance (ordonné des graphiques) en fonction du pourcentage de bruit dans l'image (abscisse des graphiques). On voit dans les graphiques que la qualité de la discrimination diminue rapidement lorsqu'on augmente le pourcentage de bruit. Pour le btr152 vu à 180 degrés, on remarque

même une fausse détection à partir de 5% de bruit. Le calcul du pourcentage de bruit se fait de la façon suivante:

$$\% \textit{ de bruit} = \frac{\text{énergie du bruit}}{\text{énergie du bruit} + \text{énergie de l'objet}} \cdot 100$$

Il y a donc deux éléments majeurs à identifier pour savoir si le bruit dégradera beaucoup la reconnaissance: l'énergie totale du bruit par rapport à l'image à reconnaître et la répartition de cette énergie dans le spectre de puissance total. Nous allons analyser tout d'abord un exemple d'effet de l'énergie du bruit en ajoutant du bruit blanc à une des images de la banque, soit le léopard vu à 90 degrés.

8.2.2 Bruit de fond

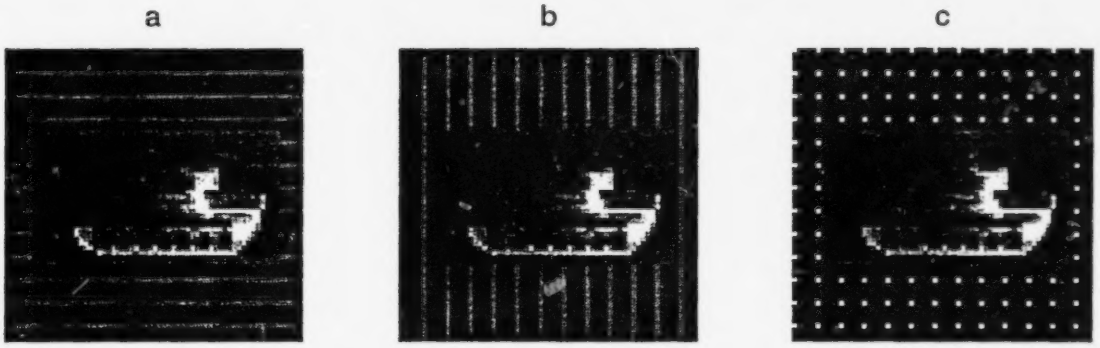


Figure 30: Les trois type de bruits de fond utilisés avec un objet. a) Bruit horizontal, b) Bruit vertical, c) Bruit carré

Pour étudier le bruit de fond, on utilise trois types de scènes: des barres horizontales, des barres verticales et un quadrillage de points. L'avantage de ces types de fonds est l'allure de leur transformée de Fourier: pour le motif à barres horizontales, on retrouve des composantes spectrales presque uniquement verticales, l'inverse se produit pour les barres verticales et pour le quadrillage de points, on retrouve une répartition sensiblement égale entre la verticale et l'horizontale (voir les vecteurs de sortie du compresseur de ces bruit de fond au Tableau 5).

Vecteurs de sortie du compresseur en secteur pour des bruits de fond								
Bruit de fond	**Vecteurs**							
fond horizontal:	0.133	0.037	0.043	0.060	0.058	0.067	0.099	0.660
	0.694	0.099	0.066	0.057	0.060	0.043	0.037	0.126
fond vertical:	0.685	0.101	0.066	0.053	0.060	0.044	0.040	0.143
	0.153	0.040	0.044	0.060	0.059	0.066	0.100	0.660
fond carré:	0.416	0.071	0.085	0.266	0.264	0.082	0.069	0.402
	0.422	0.068	0.083	0.263	0.265	0.085	0.070	0.400

Tableau 5

Les graphiques des Figure 34 à Figure 36 montrent la distance de reconnaissance du léopard vu à 90 degrés en fonction du pourcentage du bruit dans l'image. On remarque que pour ce type de bruit nous pouvons tirer les mêmes conclusions que précédement, soit que la discrimination se détériore avec l'augmentation du pourcentage de bruit. On voit aussi très bien l'effet des différences entre les répartitions d'énergie dans les vecteurs. Ainsi, pour le même objet, on remarque le début de fausses détections à 3,5% de bruit pour le fond de barres verticales, à 4% pour le fond à barres horizontales. Pour le fond de quadrillage, cette fausse détection débute à 8%.

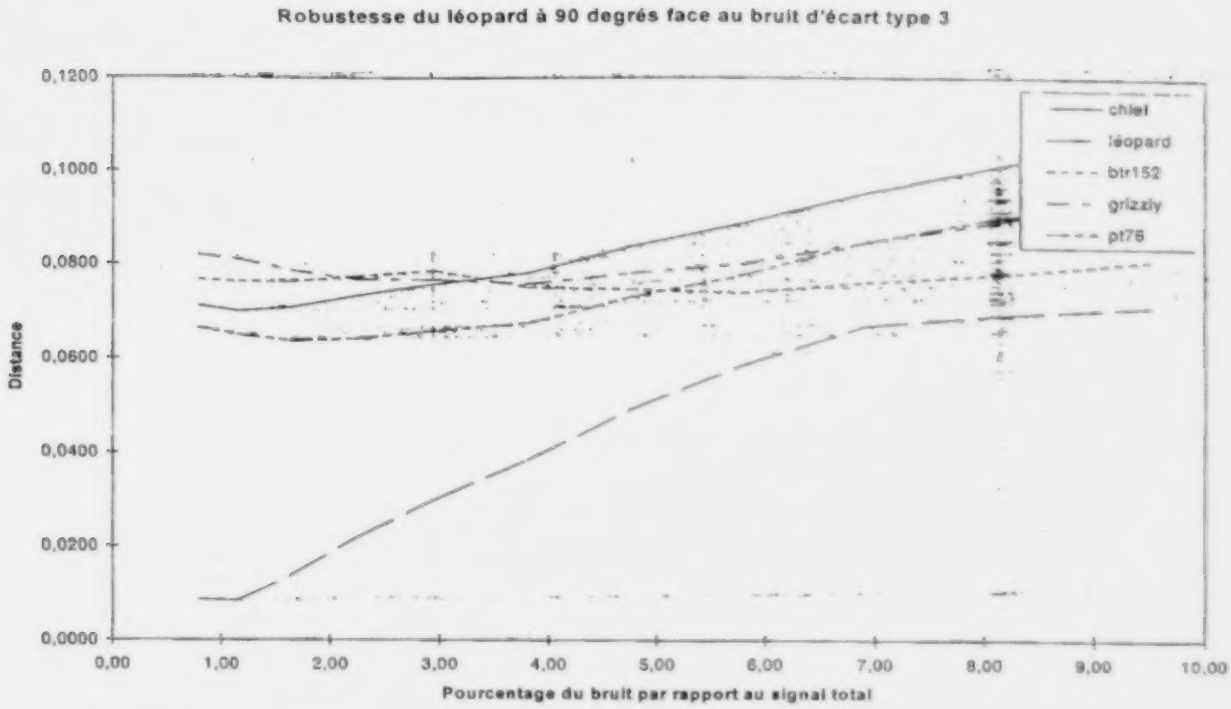


Figure 31

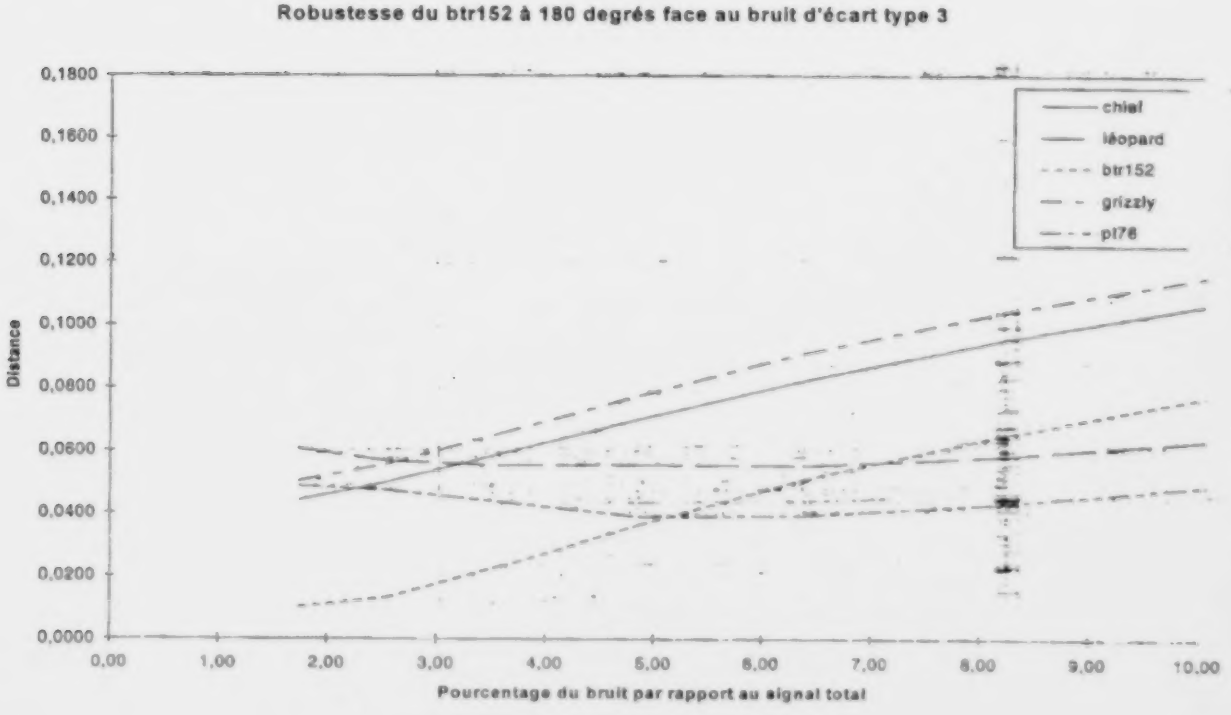


Figure 32

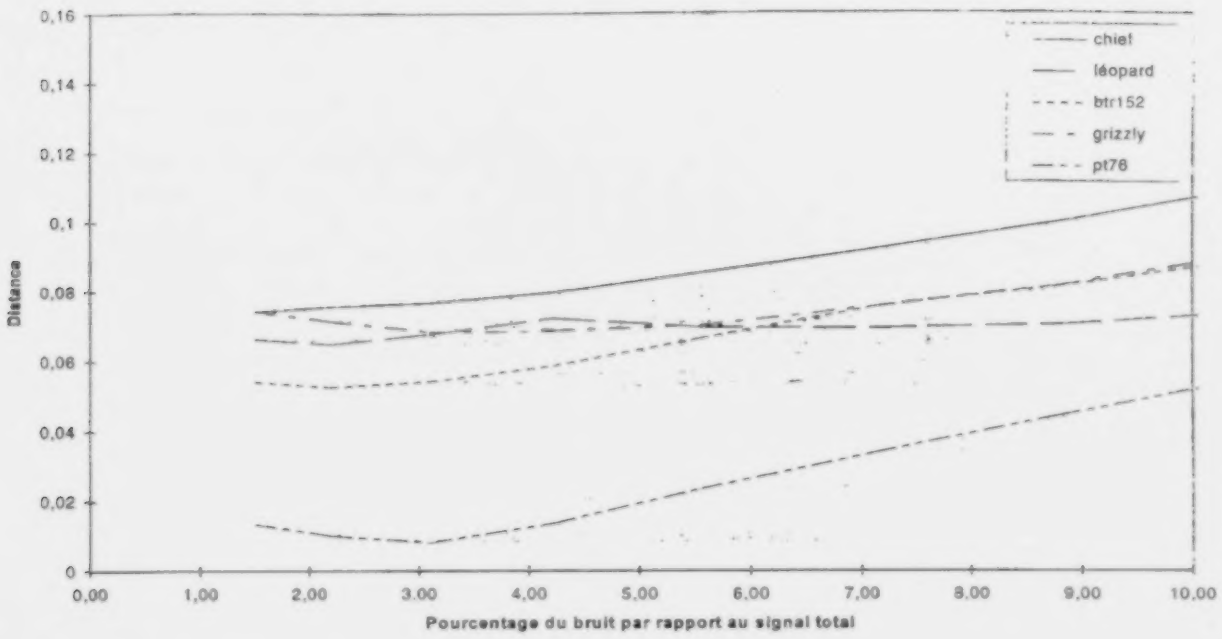

Robustesse du pt76 à 270 degrés face au bruit d'écart type 3
0,16
0,14
0,12
0,1
0,08
0,06
0,04
0,02
0
Distance
0,00
1,00
2,00
3,00
4,00
5,00
6,00
7,00
8,00
9,00
10,00
Pourcentage du bruit par rapport au signal total
chief
léopard
btr152
grizzly
pt76

Figure 33

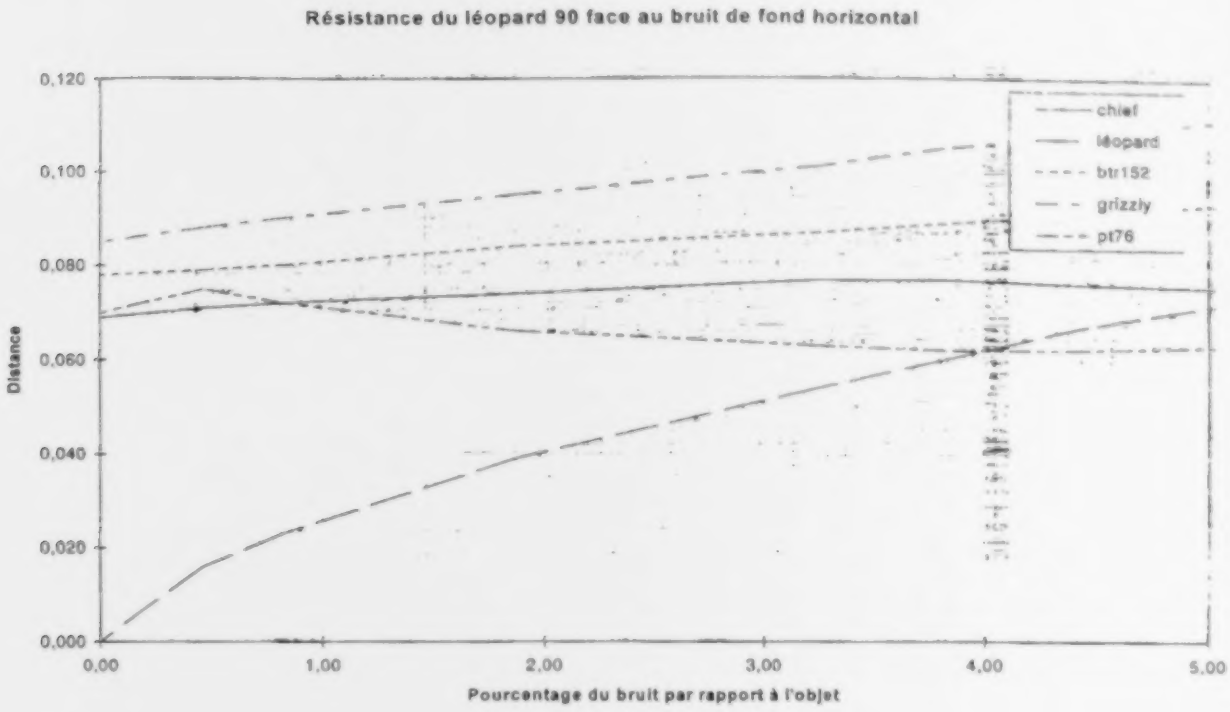


Figure 34

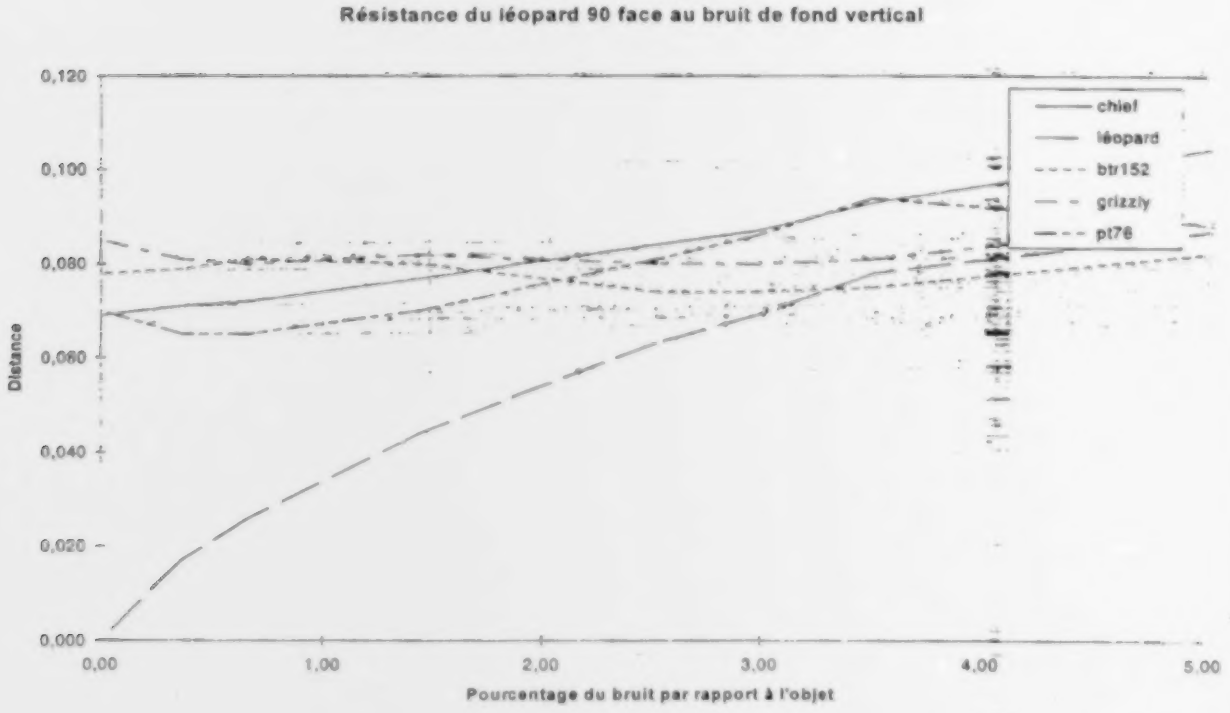


Figure 35

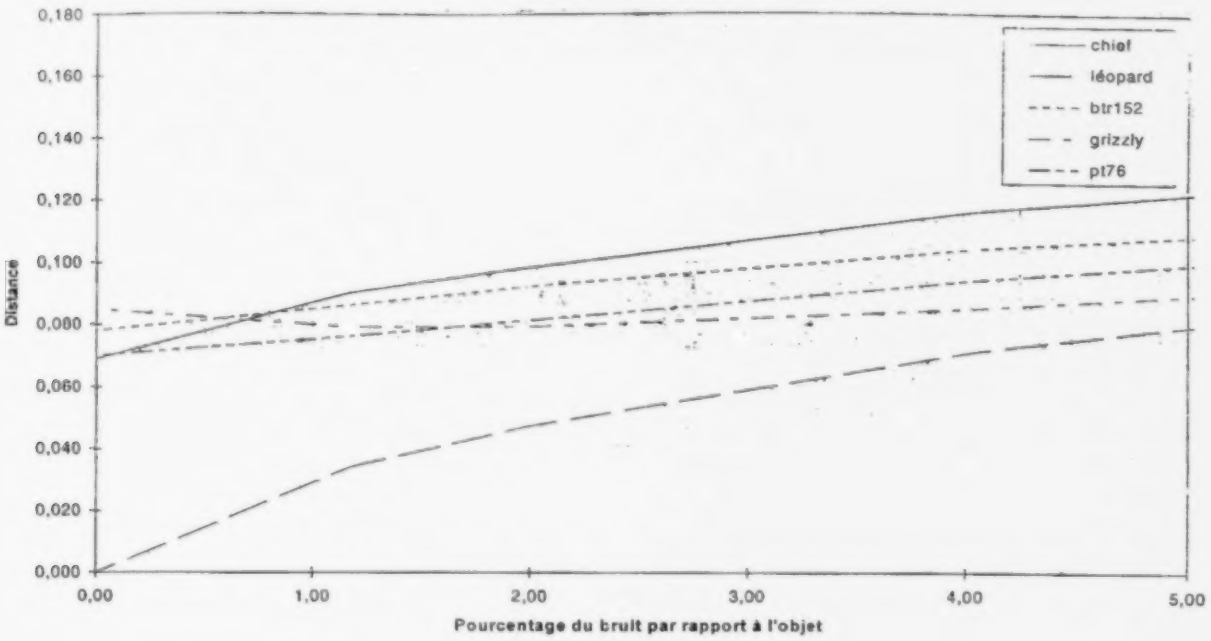

Résistance du léopard 90 face au bruit de fond carre
0,180
0,160
0,140
0,120
0,100
0,080
0,060
0,040
0,020
0,000
Distance
0,00
1,00
2,00
3,00
4,00
5,00
Pourcentage du bruit par rapport à l'objet
chief
léopard
btr152
grizzly
pt76

Figure 36

8.3. Capacité de reconnaissance

Nous allons étudier dans cette section la capacité de reconnaissance et de classification du système général avec 120 images à tous les 15 degrés en apprentissage et 240 images en test. Le système comprend le filtrage moyen et la normalisation pour s'assurer d'avoir la meilleure invariance sous échelle. Dans un premier temps nous regarderons l'effet du réseau de neurones et ensuite nous tenterons de modifier l'apprentissage pour avoir un meilleur résultat de reconnaissance et de classification.

8.3.1 Réseau de neurones

D'après Neiberg et Casasent, l'ajout d'un réseau de neurones linéaire avec comme poids la banque d'images de l'apprentissage permet d'augmenter la qualité de la reconnaissance. Le tableau 6 résume les résultats de reconnaissance et de classification avec ce type de réseau dans notre montage tels qu'illustrés dans les graphiques des Figure 37 à Figure 41.

véhicule	CHIEF	LÉOPARD	BTR152	GRIZZLY	PT76	TOTAL
mauvaise reconnaissance	3	3	8	12	9	35
mauvaise classification	2	3	1	0	4	10

Tableau 6: Analyse de la reconnaissance et de la classification avec le réseau de neurones

On trouve un taux de reconnaissance de 90.3% en tenant compte de toute la banque d'images et 85% en ne tenant pas compte des images utilisées pour l'apprentissage. Si on compare ces résultats de reconnaissance avec ceux du montage avec filtrage et normalisation de la section sur l'invariance d'échelle, 95% avec les images d'apprentissage et 92.5% sans les images d'apprentissage on trouve une différence de 7.5%, soit 17 images qui ne sont plus reconnues par le système lors de l'utilisation du réseau de neurones. Au niveau de la classification, les résultats ne sont pas aussi différents mais il y a tout de même une dégradation. On y trouve 97.2% avec les images d'apprentissage et 95.8% sans les images d'apprentissage comparativement à 97.8% avec apprentissage et 96.7 sans apprentissage pour le montage sans réseau de neurones. Il y a une différence de 0.9%, soit 2 images qui ne sont plus classifiées correctement par le système lors de l'utilisation du réseau de neurones. On remarque donc que nos résultats vont à l'encontre de ce qu'ont trouvé Neiberg et Casasent dans leurs travaux. Ceci peut être causé par le nombre deux fois plus élevé d'images d'apprentissage utilisées dans notre étude, soit 120 images au lieu de 66. On sait que pour les réseaux de neurones linéaires il est très difficile d'avoir de bonnes orthogonalisations des vecteurs lorsque la dimension des vecteurs d'apprentissage est beaucoup plus petite que le nombre de ces vecteurs d'apprentissage. Dans notre cas, la dimension des vecteurs est 16 et le nombre de vecteurs d'apprentissage est 120, soit un rapport de 7.5. De plus, en étudiant l'allure générale des vecteurs d'apprentissage, on remarque qu'ils sont très semblables d'un objet à un autre. Ceci est causé par la nature même des objets qui ont des composantes spectrales dominantes selon l'axe des x

et l'axe des y. Cela diminue encore plus l'orthogonalité initiale des objets les uns par rapport aux autres. Ainsi notre système donne de meilleurs résultats sans réseau de neurones.

8.3.2 Amélioration de l'apprentissage

Jusqu'à maintenant, tous les apprentissages se faisaient avec des images à tous les 15 degré de rotation. L'intérêt de faire un apprentissage à cet intervalle est la faible quantité d'images d'apprentissage pour caractériser avec toute la plage de rotation hors plan. On remarque cependant qu'il y a des régions où la reconnaissance se fait plus difficilement. Par exemple, la majorité des fausses reconnaissances et des fausses classifications se fait lorsque les véhicules sont vue de dos et de face (dans la région de 0 degré et dans la région de 180 degrés). Au contraire, on remarque qu'il y a des régions où la reconnaissance se fait très facilement et que certains apprentissages ne semblent pas apporter d'information utile. C'est le cas par exemple du chief à 45 degrés, du grizzly à 105 degré. Il serait donc intéressant de tenter d'optimiser un peu plus la façon dont on fait l'apprentissage pour conserver le même nombre d'images apprises tout en concentrant cet apprentissage dans les régions critiques et en l'espaçant dans les régions de reconnaissance facile.

On retrouve dans les graphiques des Figure 42 à Figure 46 cette amélioration de l'apprentissage. On peut voir que les résultats de reconnaissance et de classification sont de 100% pour la banque d'images test que nous avons. Il serait intéressant de pouvoir avoir des images intermédiaires à celle que nous avons pour vérifier la qualité de la reconnaissance continue et non plus discrète. Ainsi nous aurions la capacité de reconnaissance et la capacité de classification exacte de notre montage.

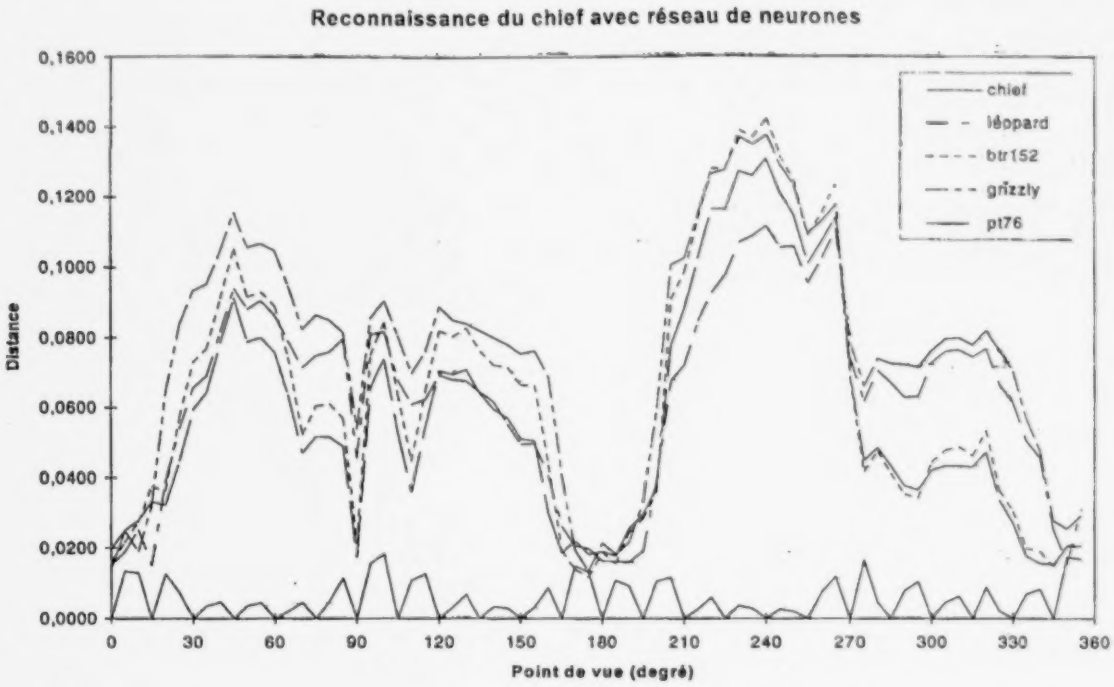


Figure 37

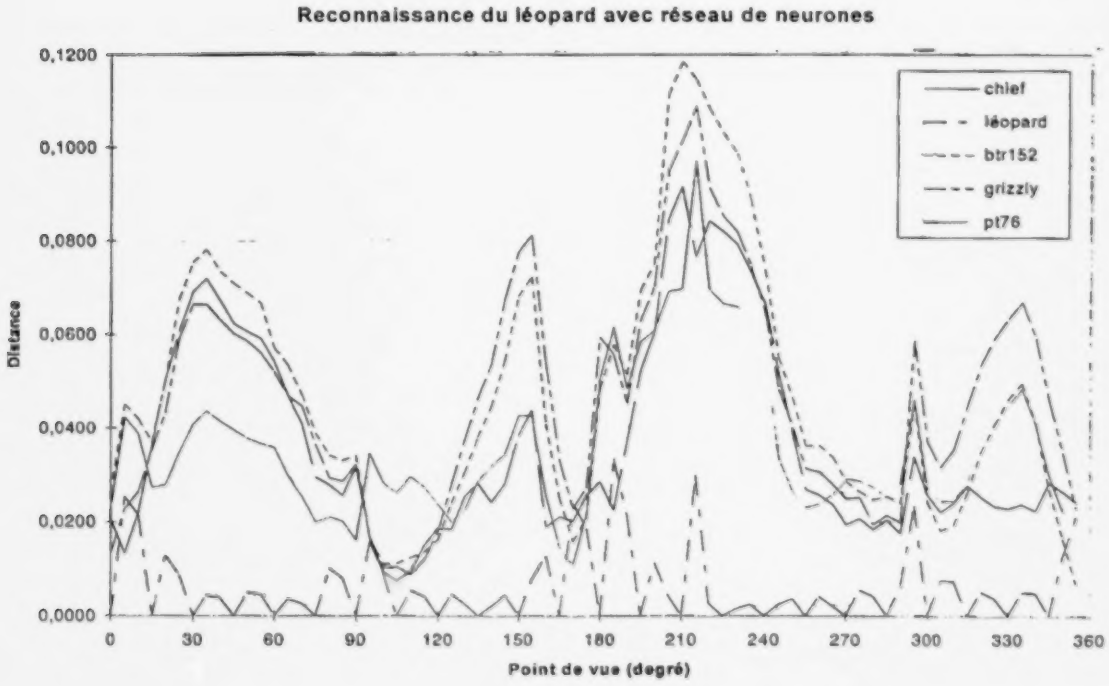


Figure 38

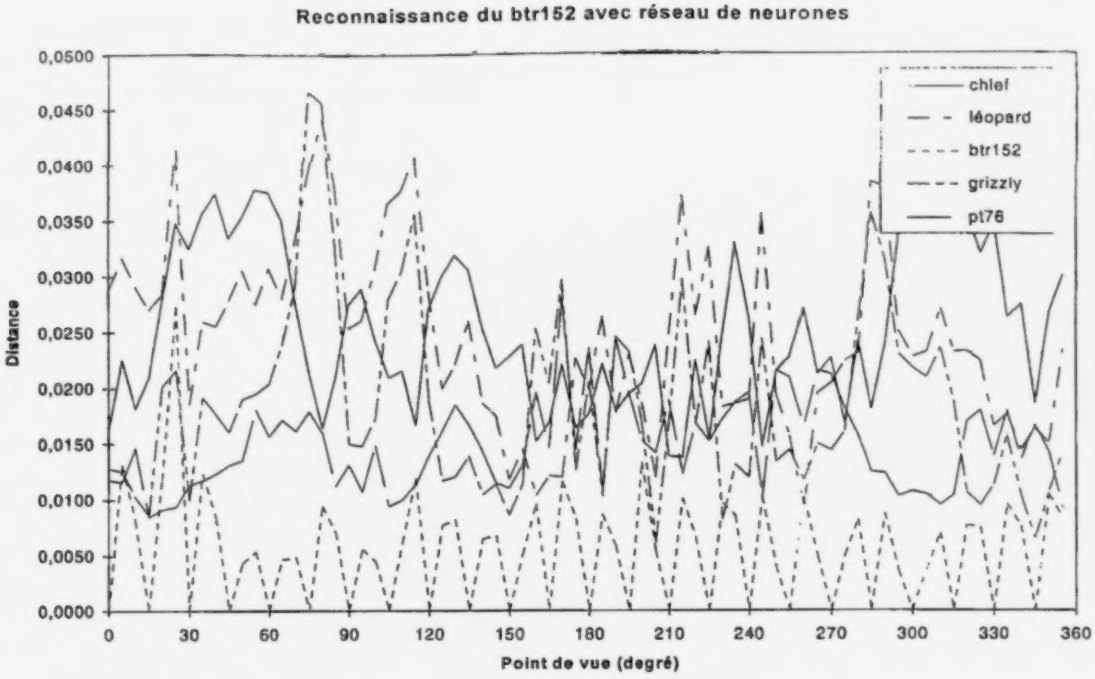


Figure 39

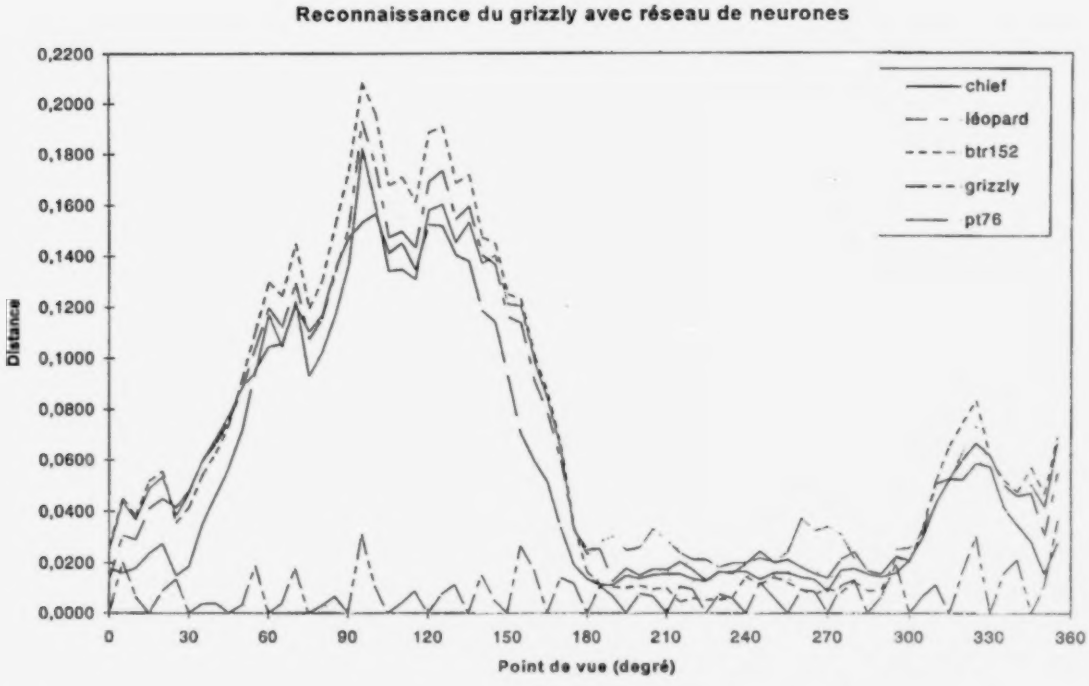


Figure 40

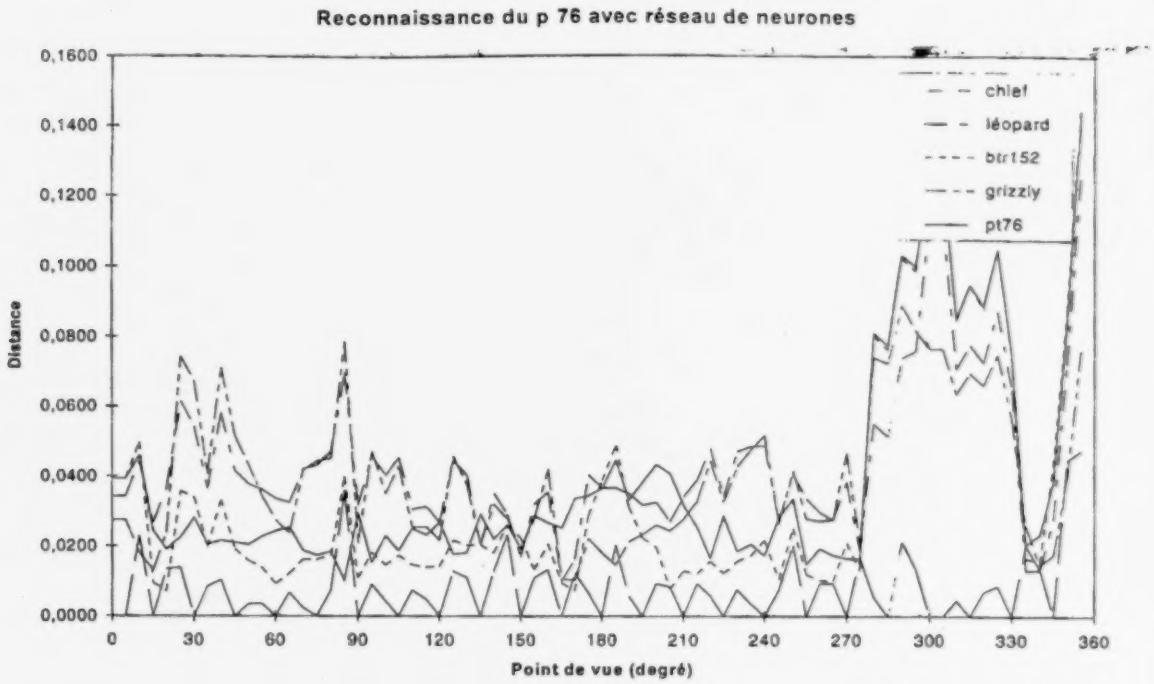

Reconnaissance du p 76 avec réseau de neurones
0,1600
0,1400
0,1200
0,1000
0,0800
0,0600
0,0400
0,0200
0,0000
Distance
0
30
60
90
120
150
180
210
240
270
300
330
360
Point de vue (degré)
chief
léopard
btr152
grizzly
pt76

Figure 41

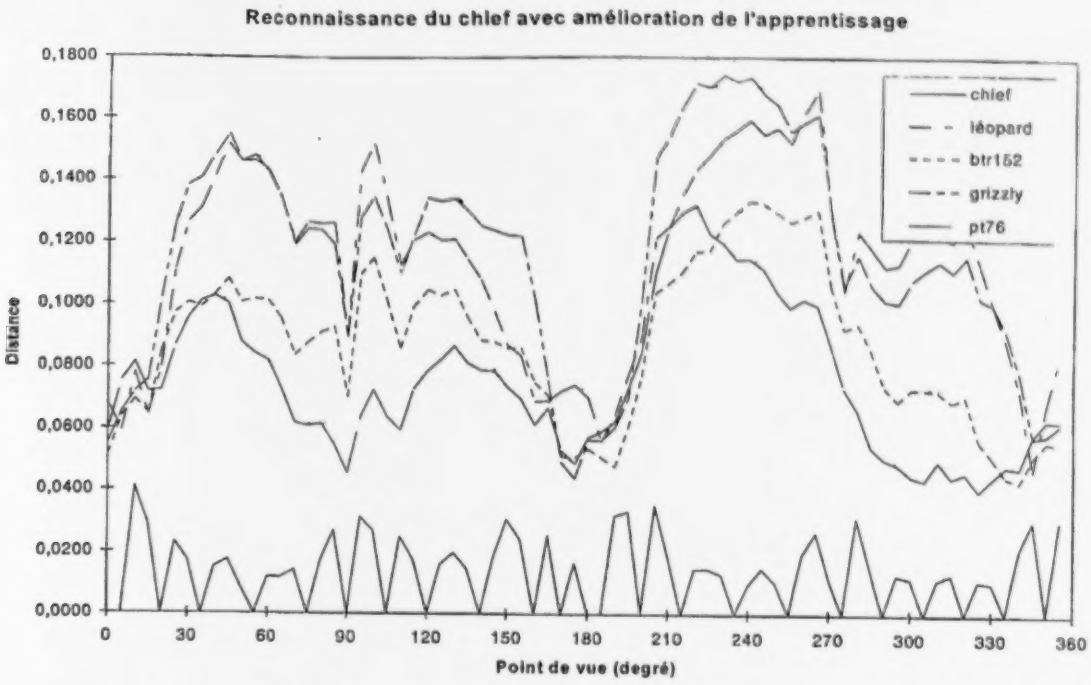


Figure 42

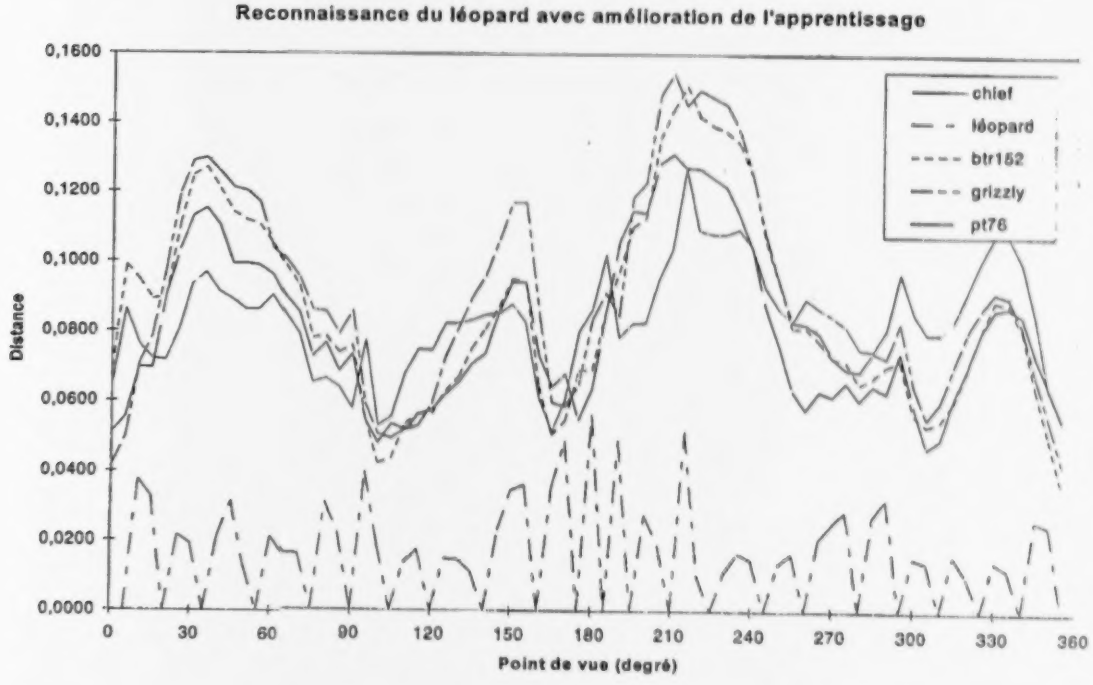


Figure 43

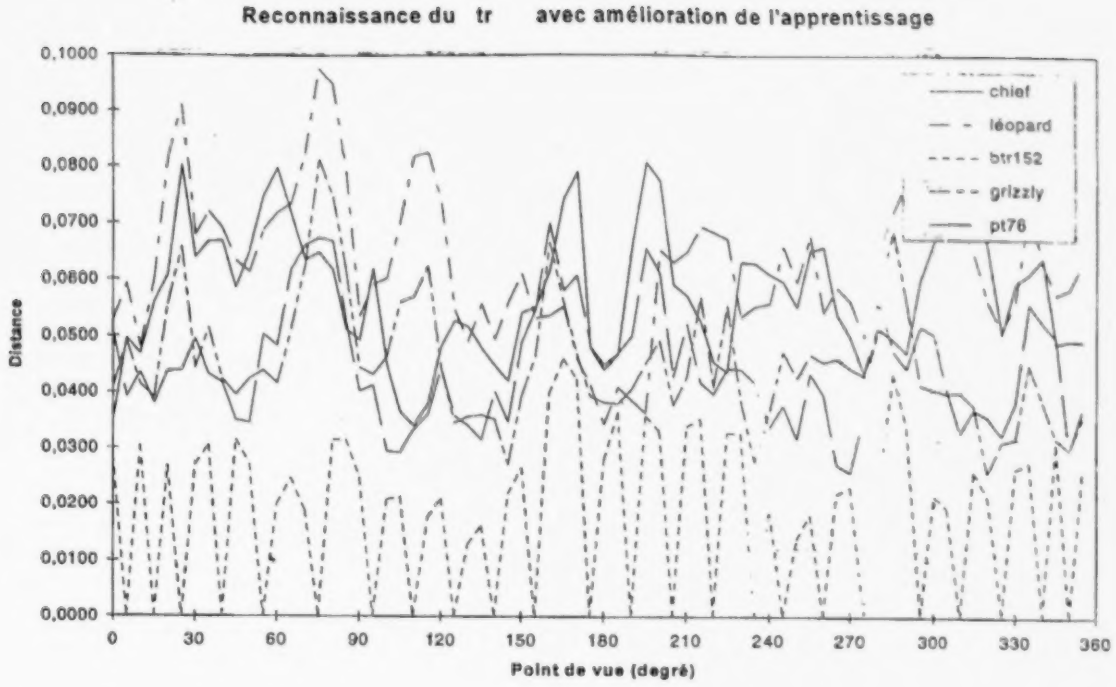


Figure 44

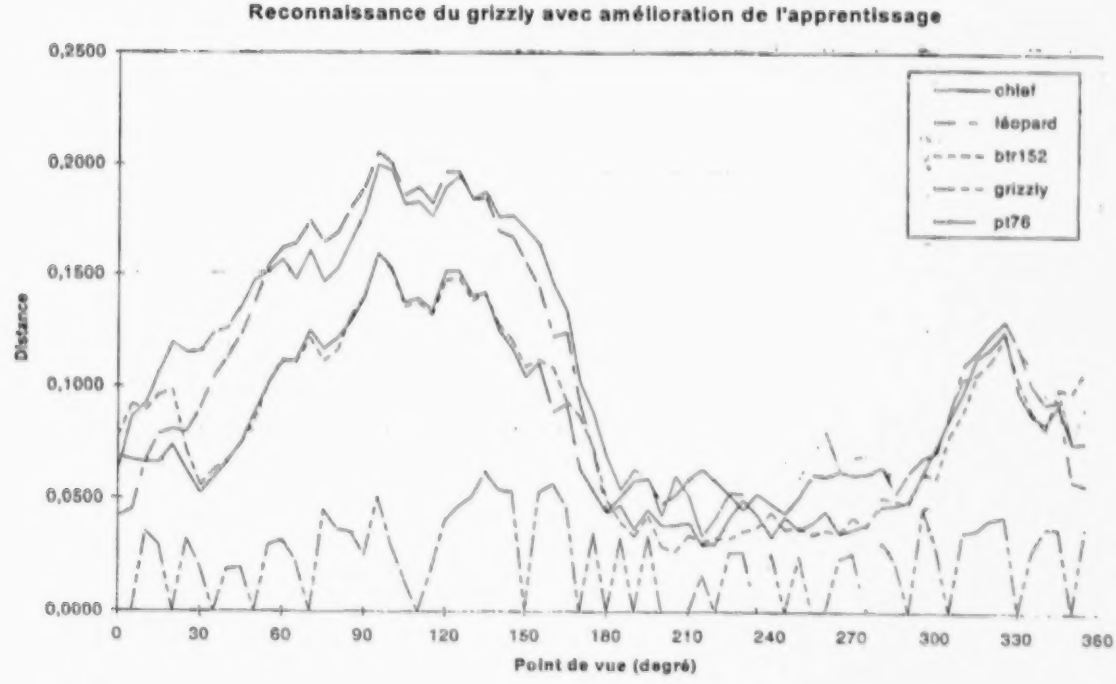


Figure 45

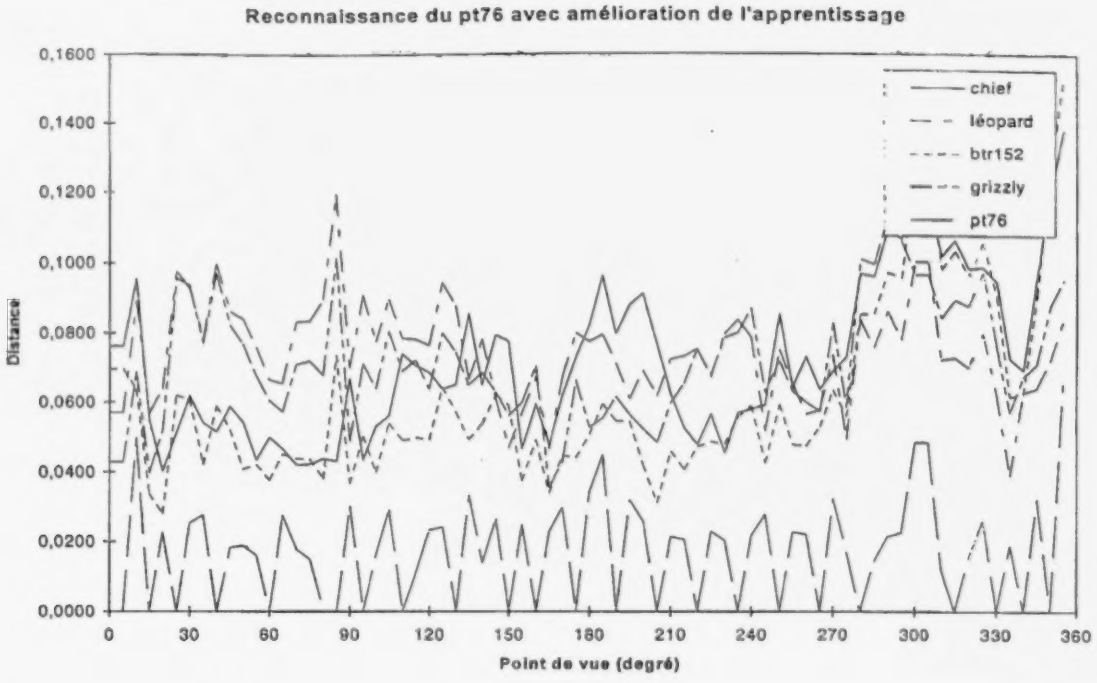

Reconnaissance du pt76 avec amélioration de l'apprentissage
0,1600
0,1400
0,1200
0,1000
0,0800
0,0600
0,0400
0,0200
0,0000
Distance
0
30
60
90
120
150
180
210
240
270
300
330
360
Point de vue (degré)
chief
léopard
btr152
grizzly
pt76

Figure 46

Deuxième méthode

9. Introduction

Le problème qui nous fut proposé par l'INO vise à étudier une technique de reconnaissance automatique de cible (RAC) dans l'espace (3-D). Les cibles sont des véhicules militaires qui circulent dans une situation réaliste sur le terrain; nous tâcherons de les détecter et de les reconnaître. Le problème qui nous est soumis peut être divisé en deux volets principaux: la reconnaissance d'un objet tridimensionnel et la RAC.

Dans l'aspect traitant de la reconnaissance de l'objet 3-D, la cible peut apparaître avec une grande variation de dimensions et d'orientations, autant dans le plan qu'en dehors du plan de l'image. Dans notre projet, les véhicules à détecter présentent une grande variation d'échelle et d'angle de perspective. Les véhicules peuvent se trouver dans toutes les orientations rencontrées lors d'une rotation de 360 degrés en dehors du plan de l'image.

Au niveau de la RAC, nous avons affaire à un environnement non-controlé où il y a une très grande variété d'objets parasites en fond de scène. La cible peut aussi être cachée par les objets parasites, la fumée en émanant de même que les reliefs accidentés du terrain lui-même. On a pu observer une telle situation lors du visionnement du vidéo, présenté par les responsables de l'INO, et montrant des véhicules militaires circulant sur le terrain. L'image de la cible peut avoir un faible contraste et être de faible résolution. Dans le cas d'images infrarouges, la luminosité varie de plus avec la condition climatique de l'environnement et avec le fonctionnement des appareils du véhicule. C'est notamment le cas lorsque le moteur et le canon sont en fonction, leurs températures augmentant. Ces parties deviennent alors plus brillantes dans les images infrarouges. Les problèmes mentionnés ci-haut représentent en eux des sujets de recherche dans le domaine de la RAC.

On voit donc que la RAC est une tâche difficile. Les aspects qu'elle considère sont très différents en comparaison avec les applications commerciales de la reconnaissance des formes, comme la reconnaissance de la parole et la vision robotique. Remplir cette tâche requiert généralement l'utilisation de plusieurs senseurs de différents types, comme les signaux radar et sonar, les images optoélectroniques multispectrales et les images stéréo (range image). Une fusion temporelle de ces données peut améliorer grandement le résultat de la classification.

Toutefois, notre projet est défini de telle sorte que nous devons procéder à la reconnaissances des véhicules à partir d'images bidimensionnelle captées par une caméra infrarouge, donc des images 2-D provenant d'un seul senseur. Nous ne possédons pas d'information sur la troisième dimension (profondeur de l'objet). Par conséquent, on ne peut assurer un taux de détection et de classification parfaite à 100% dans tous les cas. Et même si la reconnaissance était réalisée à 100% dans certaines démonstrations expérimentales, il est entendu que ce taux de réussite serait sûrement inférieur si on aborde des exemples dont l'environnement a été changé.

Pour faciliter notre travail, il fut entendu, par notre groupe et les responsables de l'INO, de considérer uniquement les images des véhicules stationnés sur un plateau tournant. Ces images ont été prises à tous les 5 degrés et on ne détecte pas d'objets parasites dans le fond. Le projet consiste alors davantage à un processus de reconnaissance d'objet tridimensionnel. Cependant, les images étant prises dans l'infrarouges, nous obtenons alors de faibles contrastes. Cette caractéristique va parfois même jusqu'à la limite où un observateur humain est incapable de percevoir les bords de l'objet. De plus, l'intensité lumineuse est non-uniforme sur l'objet. La

partie représentant le moteur de l'appareil est extrêmement brillante à cause de la température élevée dégagée durant son fonctionnement. Ces aspects correspondent évidemment à des problèmes propres à la RAC. Dans ce cas, la segmentation de l'image d'une scène, le rehaussement du contraste de l'image, la réduction du bruit et l'extraction du contour de l'objet sont des étapes clefs de la réussite. Après la réalisation de ces dernières, il ne reste qu'à effectuer la reconnaissance de l'objet 3-D, qui est constituée par la description des contours des images et la classification grâce à des réseaux de neurones.

L'approche qui fut proposée et étudiée par Yunlong Sheng et Sylvain Deschênes dans le but de réaliser les objectifs du projet se compose de la séquence d'opérations suivante. Dans un premier temps, les images de la banque de données ont subi un prétraitement visant à rehausser la qualité de l'image et à réduire le bruit présent. Ensuite, dans la deuxième étape, l'emphase fut mise sur l'extraction des contours de l'image par la transformée en ondelettes (TO). Dans le troisième volet de ce projet, le contour des images fut décrit par les descripteurs de Fourier (DF). Finalement, la reconnaissance des cibles par réseau de neurones et la classification des cibles fut effectuée.

Cette approche est basée sur les contours de l'image plutôt que sur l'image elle-même. Ce traitement sur les contours a l'avantage de préserver l'information sur la structure générale de l'objet tout en réduisant la quantité de données nécéssaires à l'analyse de l'image. Dans notre projet, l'utilisation des contours vise aussi à éviter le problème du changement de distribution d'intensité propre à l'analyse d'images infrarouges et causé par le mise en fonction de mécanismes du véhicule, changeant ainsi les températures des différentes parties de véhicule. Cette approche peut être considérée "classique" et chacune des étapes suivie peut constituer un domaine de recherche en soi. Dans le cadre de ce contrat, avec le temps et les ressources limitées, nous procédons à une brève revue des techniques principales pour chacun des domaines et réalisons ensuite une méthode appropriée au problème posé. La nouveauté de l'approche proposée réside dans l'utilisation de la transformée en ondelettes (TO) pour extraire le contour de l'objet et des descripteurs de Fourier (DF) pour une description du contour avec invariance sous rotation dans le plan, changement d'échelle et translation. Cette dernière étape correspond à l'utilisation du Néocognitron dans le système SAHTIRN[21] qui fut exposé brièvement dans la revue des méthodes non-étudiées. Les DFs sont invariants sous rotation dans le plan-image et sous changement d'échelle alors que le Néocognitron est invariant sous distorsion de l'image. Pour la classification, nous utiliserons un réseau de neurones basé sur la trajectoire dans l'espace des caractéristiques (TEC), ce qui est mieux adapté à la reconnaissance d'objets 3-D à partir de ses images 2-D.

Certaines parties de l'approche proposée, telle la transformée en ondelettes (TO) pour extraire des contours, pourront être réalisées optiquement. Par contre, d'autres parties, comme le rehaussement des images, la segmentation de la scène d'entrée et la description mathématique des contours sont difficiles à réaliser optiquement et ainsi ne profitent pas des véritables avantages de l'optique, tels que la rapidité et le paralléllisme.

Dans la première section étudiée, une méthode d'égalisation de l'histogramme est introduite pour balancer la distribution de l'intensité de l'image infrarouge. Pour éliminer les variations de gradient importantes au niveau du bruit de fond, nous procédons également à une série de filtrages médians. Ces prétraitements sont effectués dans le but de faciliter l'extraction des contours, qui seront trouvés grâce à la TO. L'avantage de cette technique par rapport aux autres algorithmes de recherche de contours réside dans son traitement à partir d'images lissées. De plus, la TO permet de procéder à une analyse multirésolutionnelle des images en traitement.

Ensuite, nous étudions les DFs, qui décrivent le contour, et les comparons aux autres descripteurs existants. Dans la quatrième section, nous donnons une justification de notre utilisation du modèle de réseau de neurones TEC par rapport aux autres modèles pour la reconnaissance d'objets 3-D. Les résultats de notre recherche montrent d'ailleurs un taux de succès élevé pour la reconnaissance avec cette approche dans les conditions de notre expérience.

10. Rehaussement des images

Comme nous l'avons vu, la reconnaissance automatique de cible (RAC) correspond à une détection de cibles dans un environnement non-controlé. Le rehaussement du contraste, la réduction du bruit, la segmentation de l'image et l'extraction des contours sont, de ce fait, des étapes-clefs de la RAC.

10.1. Image originale

Pour des images avec un bon contraste et dépourvue de bruit, l'extraction de coutour se réalise très facilement. Dans ce contrat, nous travaillons toutefois avec des images infrarouges dans un environnement réaliste. L'environnement de nos images est partiellement contrôlé, les véhicules étant stationnés sur un plateau tournant contrairement à une image prise sur le terrain. On ne retrouve pas non plus d'objets parasites dans le fond. Cependant, les images infrarouges ont typiquement de faibles contrastes et d'importants bruits qui sont principalement dus à la nature non-uniforme du gradient thermique à l'arrière-scène de l'image. Pour tester la méthode proposée, nous commençons par une image de notre banque de données qui est de meilleure qualité, avec moins de bruit, comme montrée à la Figure 47.

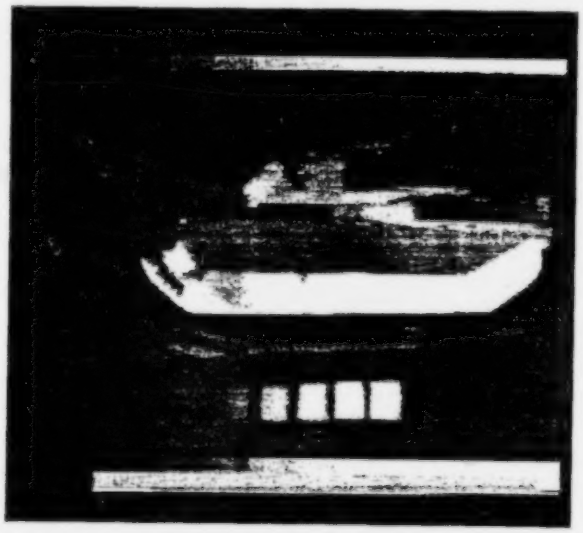


Figure 47. Image d'un char d'assaut CHIEF à un angle de 100 degrés par rapport à sa vue de face.

Dans l'image originale, la silhouette de l'objet consiste en lignes régulières: lignes droites et un petit nombre de lignes courbes. L'image du véhicule est composée de plusieurs régions de différentes intensités (température), plus ou moins uniformes à l'intérieur de la dite région. Certaines parties de l'image du véhicule ont un contraste si faible que leurs bords sont très difficiles à déterminer, même pour un observateur humain. Ce faible contraste est une conséquence directe du fait que ces parties de véhicule sont à la même températures que l'atmosphère ambiante. On remarque qu'une partie de la base du véhicule est très brillante à cause de la température élevée du moteur et des roues. Le prétraitement que nous ferons sur les

images visera à en rehausser la qualité et ainsi éliminer partiellement les problèmes qu'apportent ces caractéristiques nuisibles à l'étape d'extraction des contours.

10.2. Égalisation de l'histogramme

Un problème particulier inhérent aux images infrarouges étudiées est la très grande différence entre les intensités de différentes parties du véhicule. Ceci devient encore plus problématique dans le cas fréquent où ces variations se produisent durant le fonctionnement du véhicule. La Figure 48 montre l'histogramme de l'intensité lumineuse de l'image originale. La majorité des pixels de l'image possède une très faible luminance. La partie chaude du véhicule correspond à un pic dans la partie droite de l'histogramme. Les parties sombres et des faibles contrastes de l'image sont difficilement perceptibles, même à l'oeil. Dans ces condition, l'extraction de contours devient très difficile.

Figure 48. Histogramme du nombre de pixels en fonction de l'intensité lumineuse sur 256 niveaux de gris.

Pour faire ressortir les parties sombres de l'objet, nous avons utilisé la procédure d'égalisation de l'histogramme[30]. Le principe de cette procédure est bien connu: il redistribue tous les pixels de l'image sur les différents niveaux de gris de telle sorte que l'image ainsi traitée possède à peu près le même nombre de pixels pour chacun des 256 niveaux permis.

Soit f_j l'image originale de niveaux de gris j = 1, 2, ..., J et g_k l'image transformée de niveaux de gris k = 1, 2, ..., K. On définit une relation entre deux séries de niveaux de gris discrets $b_k = T(a_j)$ et il faut que K < J. Soient $H_F(j)$ la fraction des pixels de l'image originale qui ont un niveau de gris a_j, et $H_G(k)$, la même relation pour les pixels de l'image restituée au niveau de gris b_k. Procéder à l'égalisation de l'histogramme signifie que:

$$H_G(k) = \frac{1}{K}$$

On construit l'histogramme $H_G(k)$ à partir de $H_F(j)$ pour que les nombres cumulatifs des pixels soient égaux:

$$\sum_{n=1}^{k} H_G(n) = \sum_{m=1}^{j} H_F(m) \qquad \text{avec} \qquad b_k = T(a_j)$$

L'image traitée est obtenue à partir de l'histogramme $H_G(k)$.

La Figure 49 montre l'histogramme égalisé. On voit que les variations d'intensité trop importantes dans l'image originale sont minimisées et les parties sombres de l'objet sont mise en

évidence par ce simple procédé. Il sera possible d'encore raffiner, au besoin, la procédure d'égalisation en utilisant la modification de l'histogramme.

Figure 49. Histogramme du nombre de pixels en fonction de l'intensité lumineuse sur 256 niveaux de gris après égalisation.

10.3. Rehaussement du contraste

Nous avons donc vu que, lors de l'égalisation de l'histogramme, l'intensité lumineuse des parties sombres est bien augmentée. Notre expérience ultérieure dans la deuxième partie de cette recherche montre que cette amélioration du contraste n'est pas suffisante. A plusieurs endroits, les contrastes sont encore très faibles, faisant en sorte que la poursuite du contour nécéssite l'intervention d'une opération humaine, i.e. non-automatique. Une amélioration du contour en utilisant plusieurs techniques existentes s'impose.

Bien que nous n'ayons pas pratiqué explicitement ce genre de traitement, le rehaussement du contraste constitue lui aussi une étape-clef dans le prétraitement de l'image infrarouge. Un bon contraste de l'image est important lors de l'extraction du contour par la TO, cette technique étant sensible à la variation du contraste. De plus, nous utiliserons un seuillage dans la procédure de poursuite de contour, comme on va voir dans la deuxième étape de ce projet. Un rehaussement du contraste aurait sûrement pu nous éviter les difficultés que nous allons rencontrer dans la fermeture des contours.

10.4. Filtrage Médian

Un autre problème qui fut rencontré avec l'image de départ fut la présence abondante du bruit thermique aux alentours du véhicule. Ce bruit est causé par le détecteur infrarouge et il a des variations de gradient assez importantes, comme nous le verrons plus tard alors que nous obtiendrons une image de contours à l'aide d'un détecteur de bords. Bien que nous prévoyions une bonne capacité d'élimination du bruit en utilisant les propriétés de la transformée en ondelette, il semble plus sage d'effectuer un prétraitement par l'utilisation d'une méthode non-linéaire proposée par Tukey: le filtrage médian[30].

Cette technique consiste à remplacer la valeur en intensité de chaque pixel par la valeur médiane des pixels compris dans une fenêtre carrée centrée sur ce point d'intérêt. La fenêtre peut être de dimensions variables. Les valeurs des intensités ainsi examinées sont classées par ordre croissant. Si N valeurs sont ordonnées, pour N impair, on fixera l'intensité du pixel égale à la valeur qui a (N-1)/2 voisins d'intensité inférieure ou égale et (N-1)/2 autres d'intensité supérieure ou égale. Dans le cas où N est pair, on prendra la moyenne des deux valeurs médianes. Il est intéressant de noter que ce filtrage peut être appliqué en cascade. Il peut être appliqué à l'image déjà filtrée pour éliminer un maximum de bruit. La Figure 50 montre l'image sur laquelle on a déjà opéré une égalisation de l'histogramme et ensuite filtrée par un filtre médian de fenêtre 10x10.

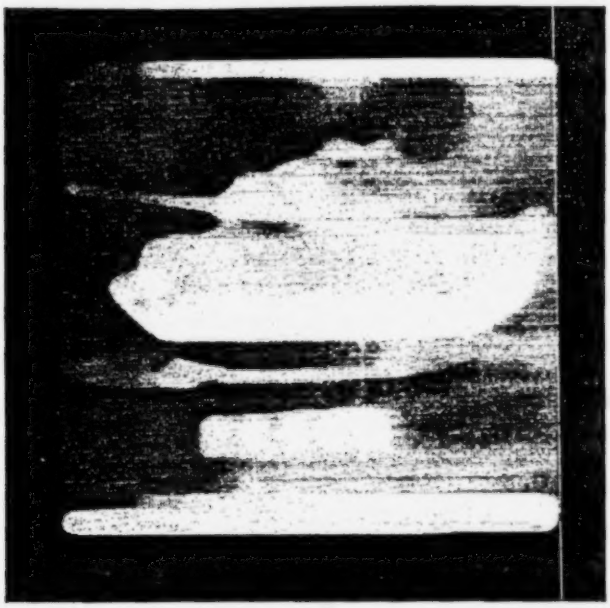


Figure 50. Image d'un char CHIEF vu à un angle de 100 degrés, préalablement traité par une égalisation de l'histogramme, après un traitement par un filtre médian de fenêtre 10 x 10.

Le filtrage médian peut cependant devenir une arme à deux tranchants: autant la technique peut éliminer le bruit indésirable, autant elle peut nous faire perdre certains signaux valides. Afin de ne pas obtenir de résultats contraires à nos attentes de départ, l'utilisation de ce filtre doit se faire avec beaucoup de circonspection. Les dimensions de la fenêtre et le nombre de filtrages sont les deux principaux paramètres nous permettant de contrôler expérimentalement le rehaussement. Nous avons pu constater, dans notre expérience, qu'une application unique d'un filtre avec une fenêtre de 10x10 nous fournit une bonne élimination du bruit sans trop grande perte de signal.

Nous avons également testé une généralisation de ce filtre, introduite par Pratt et al.[30]: le filtrage pseudomédian. Dans ce cas, la fenêtre carrée de départ est remplacée par les pixels situés sur une croix ayant toujours le pixel d'intérêt pour centre. Le principal avantage de ce nouveau traitement réside dans la conservation de signaux de bords du véhicule. L'expérience montre en effet que le filtre pseudomédian n'affecte pas l'image binaire d'un carré comme il est montré dans la Figure 51. Cependant, les tests effectués n'ont pas montré de grande amélioration apportée par le pseudo-médian et pour le reste de notre projet, nous traiterons les images avec une cascade de trois filtres médians 10x10, ce qui s'est avéré être le plus efficaces pour les traitements ultérieurs que nous appliquerons aux images.

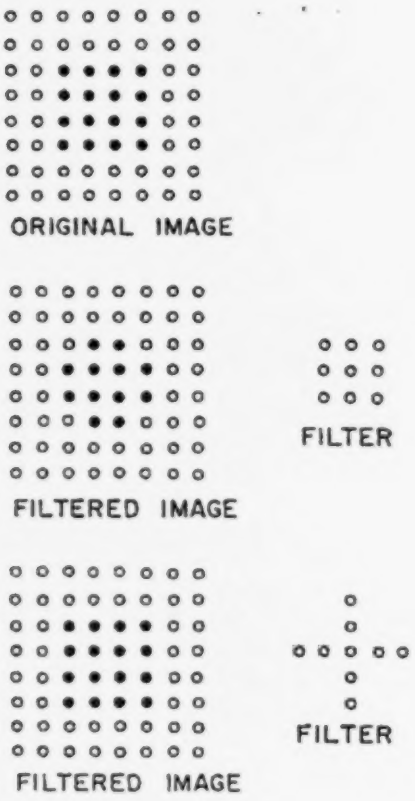


Figure 51. Effet des filtrages médian et pseudo-médian cruciforme appliqués sur un carré de 4 pixels carré.

11. Extraction du contour avec la transformée en ondelettes

L'extraction de caractéristiques des images est l'étape-clef pour réussir la reconnaissance des formes. Si l'être humain saisit facilement les caractéristiques les plus importantes d'un objet, les machines, elles, ne peuvent pas en faire autant. Le contour est l'une des caractéristiques les plus importantes de l'objet. Il contient une information beaucoup plus précise que l'intensité du spectre de Fourier de l'objet, par exemple.

Dans le cas du traitement de l'image infrarouge, le contour ne varie pas avec un changement de la température de l'objet ou de certaines de ses parties. Or, les différentes composantes (moteur, canon, etc.) d'un véhicule militaire changent de température en cours de fonctionnement. Ces changements ne sont généralement pas considérés comme des caractéristiques discriminatoires du véhicule. C'est là que réside le principal avantage de la reconnaissance basée sur le contour.

Il existe présentement deux types d'extracteurs de bords. L'un est basé sur les réseaux de neurones[31] , tandis que l'autre est de type conventionnel. L'étude montre que la méthode de Canny, qui correspond au deuxième type de méthode, a la meilleure performance[32]. Dans notre approche, nous utilisons la transformée en ondelettes pour extraire le contour.

11.1. Méthodes conventionnelles

11.1.1 Filtre de Prewitt, Sobel et Frei-Chen

Plusieurs filtres numériques existent pour la détection des bords. Les opérateurs de première dérivée calculent les gradients dans les deux directions orthogonales. Dans le cas typique des filtres 3x3, on calcule les gradients dans les directions x et y de la façon suivante:

$$G_x(j,k) = \frac{1}{K+2}[(A_2 + KA_3 + A_4) - (A_0 + KA_7 + A_6)]$$

$$G_y(j,k) = \frac{1}{K+2}[(A_0 + KA_1 + A_2) - (A_6 + KA_5 + A_4)]$$

où les A_i sont les intensités de l'image aux pixels voisins définis dans la Figure 52. Le gradient de la racine carrée est donc obtenu par:

$$G(j,k) = \{[G_x(j,k)]^2 + [G_y(j,k)]^2\}^{1/2}$$

Le détecteur de bord est de type Prewitt quand K = 1, de type Sobel quand K = 2 et de type Frei-Chen quand K = $\sqrt{2}$. Les filtres de dimension 3x3 sont rapides à calculer mais ils sont aussi plus sensibles au bruit. Les filtres de dimension plus large offrent une plus grande robustesse.

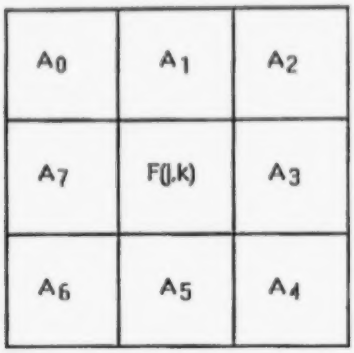


Figure 52. Intensité des pixels voisins d'un point d'intérêt défini par F(j,k) dans une fenêtre 3x3

11.1.2 Détecteur de bords optimal

Les filtres proposés par Canny[32] sont obtenus à partir d'une optimisation numérique. Elle prend en considération les différents types de bords (marche, faîte et "ridge", voir

Figure 53). Les critères à optimiser sont le rapport signal sur bruit, quand l'image est dégradée par le bruit aléatoire additif, la localisation du bord et l'élimination des pics multiples aux alentours du bord, qui correspondent à du bruit dans les bords détectés. Canny a découvert qu'un filtre de large dimension donne un meilleur rapport signal sur bruit tandis qu'un filtre étroit donne une meilleure localisation. Il y a donc un principe d'incertitude qui s'applique à ces deux critères. Puisque c'est le produit de la localisation et du rapport signal sur bruit qui est indépendant de l'échelle du filtre, il y a alors un compromis à faire entre ces deux critères. Le type de performance qu'on veut obtenir du filtre dépend donc de son échelle.

Figure 53. Types de bords considérés par l'approche de Canny lors de l'optimisation de sa méthode.

Canny propose une approche heuristique pour la détection de bords avec les filtres à échelles multiples. On prend tous les bords détectés par le filtre de plus petite échelle. Ensuite, on fait la convolution des bords avec une fonction gaussienne de la prochaine échelle supérieure dans la direction normale aux bords. Les nouveaux bords sont alors retenus. Le processus continue aux échelles plus grandes.

Pour les bords de marche, l'optimisation de Canny donne une série de filtres optimaux qui dépendent des paramètres déterminés par le compromis fait dans le choix des critères désirés. Un des filtres optimaux, donnant la capacité maximale contre le bruit formé par les multiples pics près des bords, peut être approximé par un filtre de première dérivée d'une gaussienne avec une

réduction de performance de 10 à 20%. Le filtre de première dérivée d'une gaussienne est approximativement la forme optimale pour la détection de bords. A partir du principe d'associativité, la convolution entre le filtre G_n et l'image I, est égale au gradient de la convolution entre la gaussienne G et l'image I, donné par:

$$|G_n * I| = |\Delta(G * I)|$$

où

$$G_n = \frac{\partial G}{\partial n} = n \bullet \Delta G$$

La convolution G * I est l'image lissée par la gaussienne. Cette opération de lissage est nécessaire pour réduire le bruit de l'image avant de calculer la dérivée. Les opérateurs de détection de bords utilisant les filtres numériques décrit auparavant sont les filtres locaux de dimension limitée, qui correspondent aux filtres tronqués par une fenêtre. Quand la fenêtre est un simple carré 3x3, le phénomène de Gibbs dans le domaine de la transformée de Fourier va apparaître. Comme conséquence, le bord va se déplacer et deviendra discontinu. L'opération montrée dans l'équation ci-haut correspond à calculer le gradient avec un filtre infiniment large mais d'une image dans une fenêtre définie par la gaussienne. Cette dernière est une fenêtre idéale, qui a un maximum de régularité. Le produit entre la largeur spatiale et la bande passante en fréquence de la gaussienne est égal au minimum défini par le principe d'incertitude[33].

La détecteur de bord de première dérivée d'une gaussienne cherche les maxima de la valeur absolue des premières dérivées de l'image lissée. La dérivée seconde de la gaussienne a été proposée par Marr comme détecteur de bords. Ce filtre est l'opérateur Laplacien[34], et donne un "zero-crossing" à la position d'un bord. Ces deux filtres sont presque identiques pour réaliser la détection des bords, mais la méthode de "zero-crossing" indique seulement leur position, tandis que le filtre de première dérivée contient encore l'information d'amplitude de fluctuation des bords. Les minima de la valeur absolue des premières dérivées sont également les points d'inflexion. Le filtre de dérivée seconde donne aussi un "zero-crossing" pour ces minima qui correspondent aux variations lentes de la fonction, mais pas nécéssairement aux bords de l'images.

Le détecteur de Canny nécéssite un seuillage à la sortie. Cette opération laisse le détecteur de Canny sensible aux contrastes locaux de l'image et rend la poursuite du contours difficile. Canny a proposé un seuillage avec hystérésis pour éviter les discontinuités du contour. Dans notre réalisation de l'extraction du contour avec la transformée en ondelettes, discutée dans la partie suivante, nous allons rencontrer les mêmes problèmes que ceux traités par Canny : le seuillage à la sortie, la poursuite du contour et les filtres multirésolutionnels.

11.2. Transformée en ondelette

Les détecteurs de bords optimaux de Canny sont difficiles à calculer. Cependant, ses approximations, fondées sur les dérivées de premier et deuxième ordre d'une gaussienne, satisfont aux conditions d'admissibilité de l'ondelette. En fait, ces opérations sont équivalentes à la transformée en ondelettes (TO) continue. La TO avec ces deux types d'ondelettes peut être réalisée par un corrélateur optique[35].

Nous adoptons la TO discrète et orthonormée pour effectuer la détection des bords de l'objet. La raison qui nous mène à utiliser cette TO est qu'elle est calculée avec un algorithme d'arbre qui est beaucoup plus rapide que la TO continue ou que le filtre optimal de Canny.

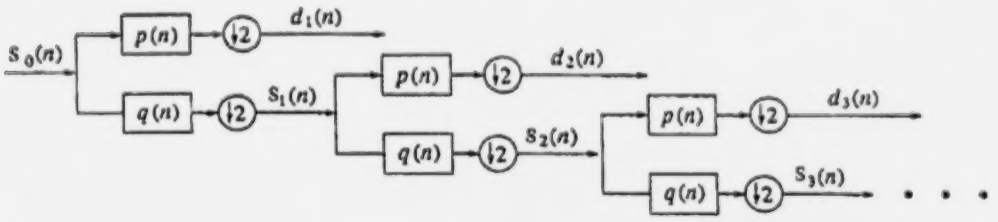


Figure 54. Algorithme d'arbre représentant la transformée en ondelette numérique à plusieurs résolutions.

La TO est une opération locale et multirésolutionnelle. La procédure de la TO numérique se fait dans un algorithme d'arbre[36] comme montré à la Figure 54. Un signal uni-dimensionnel $s_0(n)$ subit le filtrage passe-bas p(n) et le filtrage passe-haut q(n). Tous les deux sont des filtres discrets et de support fini. Dans la terminologie de la théorie de codage sous-bande (subband coding), les filtres p(n) et q(n) sont appelés filtres de miroir quadratique (quadrature mirror filters) à cause de relation entre leurs transformées de Fourier, P(u) et Q(u):

$$|Q(\tfrac{\pi}{2}-\omega)| = |P(\tfrac{\pi}{2}+\omega)|$$

Le filtrage avec p(n) donne une approximation du signal et le filtrage avec q(n) donne la différence entre le signal approximatif et le signal original. Le signal approximatif, lissé par le filtre p(n), peut être décrit avec une résolution deux fois plus faible. Un sous-échantillonnage par un facteur deux s'applique, ce qui donne $s_1(n)$. A l'étape suivante, les filtres p(n) et q(n) sont appliqués au signal approximatif et sous-échantillonné $s_1(n)$. La procédure continue jusqu'à la résolution la plus basse. Le signal d'approximation à ce niveau de résolution contient seulement deux éléments et ne porte aucune information utile. Les signaux de différence $d_i(n)$, générés par le filtre q(n) et le sous-échantillonnage à chacun des niveaux de résolution, constituent un développement d'ondelette de l'image[33].

11.2.1 Réalisation de la transformée en ondelette numérique

11.2.1.1 Transformée en ondelette unidimensionnelle

Dans le paragraphe précédent, nous avons abordé la question de la transformée en ondelette discrète dans le cas unidimensionnel. De façon à construire un algorithme, il est pertinent de la représenter comme un produit vecteur-matrice, où les éléments de la matrice sont données par les

coefficients des filtres choisis. Si on considère des filtres avec un maximum de 4 coefficients chacuns, nous obtenons une relation du genre:

$$\begin{pmatrix} p_{-1} & p_0 & p_1 & p_2 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ q_{-1} & q_0 & q_1 & q_2 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{-1} & p_0 & p_1 & p_2 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & q_{-1} & q_0 & q_1 & q_2 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & p_{-1} & p_0 & p_1 & p_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & q_{-1} & q_0 & q_1 & q_2 \\ p_1 & p_2 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & p_{-1} & p_0 \\ q_1 & q_2 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & q_{-1} & q_0 \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ y_{N-3} \\ y_{N-2} \\ y_{N-1} \\ y_N \end{pmatrix} = \begin{pmatrix} S_1 \\ D_1 \\ S_2 \\ D_2 \\ \vdots \\ S_{\frac{N}{2}-1} \\ D_{\frac{N}{2}-1} \\ S_{\frac{N}{2}} \\ D_{\frac{N}{2}} \end{pmatrix}$$

où y_i avec i = 1, 2, ..., N est le vecteur "signal", qui est multiplié par une matrice de TO. Chaque ligne dans la matrice de TO est un filtre.

On remarque que la multiplication entre le vecteur de signal et la matrice de TO donne une alternance de valeurs S et D. Les S décrivent le signal approximé alors que les D sont les détails qui représentent la différence entre ce signal approximé et le signal original. On note que la dimension des vecteurs est réduite de moitié. Ceci est dû au fait qu'en passant à une résolution plus faible de notre image lors du processus d'approximation, nous pouvons utilisé le théorème de l'échantillonnage pour diminuer le nombre de données décrivant le signal sans perdre d'information.

Dans l'opération de la matrice de TO, le signal approximé S défini par (S_1, S_2, ...) est une convolution entre le signal de départ y et le filtre passe-bas p(n) donné par (p_{-1}, p_0, p_1, p_2). De son côté, le signal de détails D défini par (D_1, D_2, ...) est la convolution entre le signal de départ y et le filtre passe-haut q(n) donné par (q_{-1}, q_0, q_1, q_2). On a donc:

$$S_k = \sum_{i=1} y_i \, p(i-2k)$$

$$D_k = \sum_{i=1} y_i \, q(i-2k)$$

Pour continuer le processus de façon à exploiter au maximum l'analyse multirésolutionnelle, il nous suffit de créer deux nouveaux vecteurs avec les valeurs S_i et D_i, i = 1, 2, ..., N/2, qu'on retrouve dans le vecteur du côté droit de l'équation. Le nouveau vecteur D, qui correspond à la TO à la première résolution, sera conservé tel quel alors que le vecteur S servira de nouveau vecteur d'entrée y sur lequel on appliquera de nouveau l'algorithme, mais en prenant une dimension réduite d'un facteur 2. Ce processus pourra être répété aussi longtemps que la

longueur du filtre choisi le permettra. Dans notre application, il est évident qu'il sera inutile de passer à des résolutions extrêmement petites, les détails du contours devenant inapplicables aux étapes ultérieures

11.2.1.2 Transformée en ondelettes d'une image bidimensionnelle

La généralisation aux problèmes bidimensionnels, comme c'est le cas dans le traitement des images digitales, est directe. Pratiquement, nous n'avons qu'à appliquer les filtres passe-haut et passe-bas aux lignes de pixels de notre image pour ensuite soumettre au même filtrage le résultat obtenu, mais en l'appliquant maintenant aux colonnes. On opère ensuite les permutations et réductions d'échelle prescrites et nous pouvons appliquer notre filtrage en cascade.

Vu la réduction en échelle inhérente à ce traitement, le résultat obtenu à partir d'une image 512x512 nous procure les quatre fenêtres montrées à la Figure 55. Dans le coin supérieur gauche, nous retrouvons l'image approximée et sous-échantillonnée (256x256) de l'image de départ. Dans la portion inférieure droite, on obtient la TO bi-dimensionnelle de l'image. Les deux derniers quadrants représentent les TO's unidimensionnelles verticale (supérieur droit) et horizontale (inférieur gauche). Comme nous le verrons par la suite, ce sont ces deux derniers quadrants qui serviront à l'extraction des bords de nos objets à reconnaître. Bien sûr, il est possible de répéter l'algorithme pour l'image approximée. On obtient alors une cascade des coefficients de la TO à plusieurs résolution comme montré à la Figure 56.

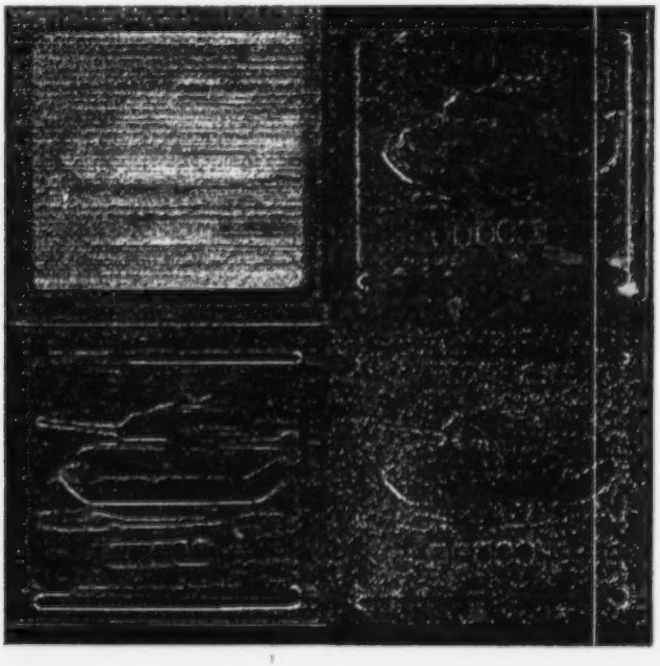


Figure 55. Résultat de l'algorithme de transformée en ondelette pour une étape de filtrage avec l'ondelette de spline quadratique.

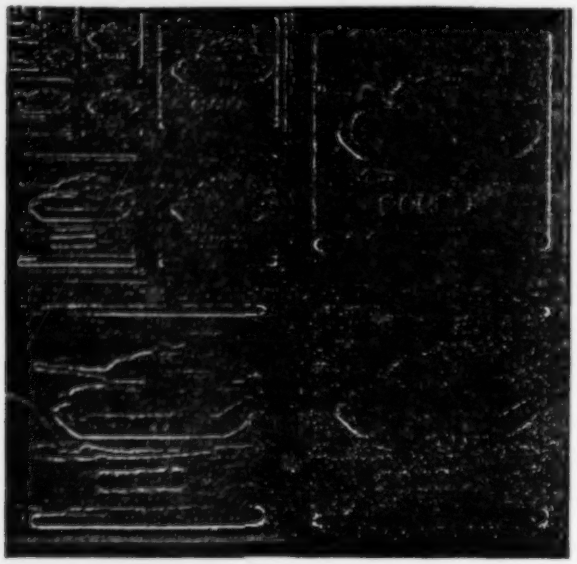


Figure 56. Résultat de l'algorithme de transformée en ondelette pour sept étapes de filtrage avec l'ondelette de spline quadratique.

Pour obtenir ces résultats, nous avons une certaine liberté au niveau du choix de l'ondelette. La TO orthonormée et discrète est effectuée dans le cadre de l'analyse multirésolutionnelle qui est basée sur une fonction d'échelle continue qui effectue un lissage. Les filtres discrets passe-bas, p(n), et passe-haut, q(n), sont construits à partir de cette fonction d'échelle. Les filtres P(ω) et Q(ω), dans le domaine des fréquences doivent de leur côté satisfaire les conditions d'orthonormalité et de régularité.

La propriété d'orthonormalité est bien connue dans la théorie des codages sous-bande. Quand les filtres P(ω) et Q(ω) dans le domaine de fréquence satisfont la condition de paraunitarité, les propriétés d'orthonormalité et de reconstruction parfaite seront satisfaites. A partir de l'orthonormalité des filtres p(n) et q(n), le filtre passe-haut peut être déduit du filtre passe-bas p(n) par la relation:

$$q(n) = (-1)^{n+1} p(N-1-n)$$

pour un nombre N pair arbitraire. Les p(n) et q(n) sont orthonormaux entre eux et ils sont également orthonormaux à leurs propres translations de 2, comme nous allons le voir plus tard dans la matrice représentant la TO de Daubechies. C'est pourquoi un nombre pair N arbitraire peut être ajouté dans la relation faisant intervenir p(N-1-n).

Les ondelettes qui ont retenu davantage notre attention sont celles de Daubechies et les splines quadratiques (Battle-Lemarié). Les ondelettes de Daubechies offrent un maximum de régularité pour une longueur donnée de filtres. Elles sont orthogonales mais ne sont pas symétriques. Ceci peut introduire des déplacements spatiaux des bords. Les ondelettes symétriques sont alors préférables pour détecter les bords de l'image. La longueur des filtres pour Daubechies peut jouer un rôle important car un filtre long sera moins sensible aux détails minimes, comme le bruit thermique. Voici les différents tests que nous avons effectués.

11.2.1.3 Ondelette de Daubechies à quatre coefficients.

Pour cette ondelette, le filtre passe-bas p(n) est donné par les quatre coefficients ($p_{-1} = c0$, $p_0 = c1$, $p_1 = c_2$, $p_2 = c_3$) et celui de passe-haut q(n) est donné par ($q_{-1} = c_3$, $q_0 = -c_2$, $q_1 = c_1$, $q_2 = -c_0$), où les c_i ont pour valeur :

$$c_0 = (1+\sqrt{3})/4 \qquad c_1 = (3+\sqrt{3})/4$$

$$c_2 = (3-\sqrt{3})/4 \qquad c_3 = (1-\sqrt{3})/4$$

On voit alors que:

$$c_0^2 + c_1^2 + c_2^2 + c_3^2 = 1$$

et aussi:

$$c_3 - c_2 + c_1 - c_0 = 0$$

$$c_2 c_0 + c_3 c_1 = 0$$

On peut vérifier l'orthogonalité entre les lignes et les colonnes dans la matrice de TO. La matrice de TO est donc une matrice orthogonale et son inverse est égale à sa transposée. Cette relation est utilisée pour la TO inverse, qui restitue le vecteur de signal y_i à partir des coefficients de la TO. Bien entendu, lors de la reconstitution que nous avons effectuée, nous n'avons obtenu aucune erreur entre l'image recontruite et celle de départ. A la Figure 57, nous présentons les résultats pour l'extraction de la TO en cascade avec l'ondelette de Daubechies à quatre coefficients.

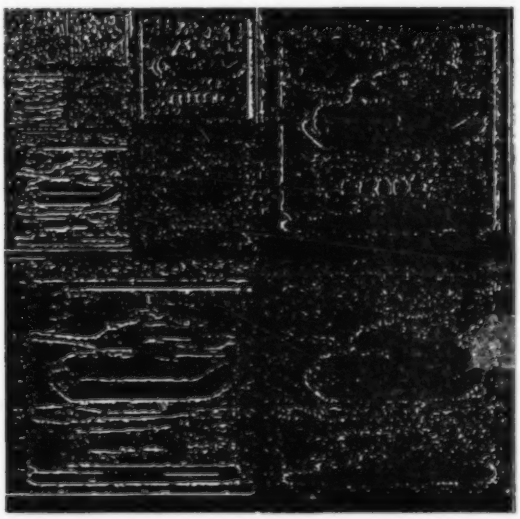


Figure 57. Résultat de l'algorithme de transformée en ondelette pour sept étapes de filtrage avec l'ondelette de Daubechies à quatre coefficients.

11.2.1.4 Ondelette de Daubechies à douze coefficients.

Nous avons repris le même algorithme en le généralisant à douze coefficients. Dans le cas de Daubechies, les coefficients pour le filtre passe-bas furent données par:

$p_{-1} = 0{,}111541$ $p_5 = 0.097502$

$p_0 = 0{,}494624$ $p_6 = 0.027523$

$p_1 = 0{,}751134$ $p_7 = -0.031582$

$p_2 = 0{,}315250$ $p_8 = 0.000554$

$p_3 = -0{,}226265$ $p_9 = 0.004777$

$p_4 = -0{,}129767$ $p_{10} = -0.001077$

Comme auparavant, nous avons défini le filtre passe-haut q(n) comme ($q_{-1} = p_{10}$, $q_0 = -p_9$, $q_1 = p_8$, $q_2 = -p_7$, ..., $q_{10} = -p_{-1}$). Nous avons finalement calculé la TO en cascade et nous sommes arrivés au résultat illustré à la Figure 58.

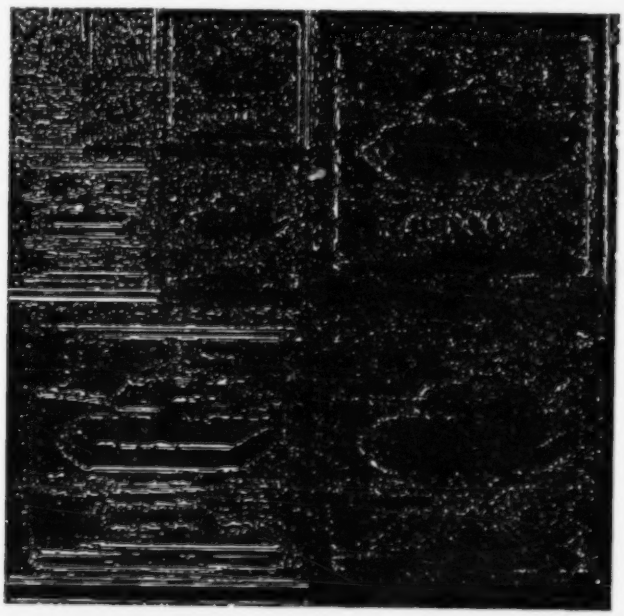


Figure 58. Résultat de l'algorithme de transformée en ondelette pour sept étapes de filtrage avec l'ondelette de Daubechies à douze coefficients.

11.2.1.5 Ondelette de spline quadratique.

Nous avons finalement opté pour l'utilisation de l'ondelette de spline quadratique, proposée par Mallat[37], pour effectuer l'extraction des contours de l'objet. Notre but est d'avoir un détecteur de

bords optimal, similaire à la première dérivée d'une gaussienne, mais pouvant être développé comme un algorithme d'arbre rapide avec des filtres numériques dans le cadre d'une analyse multirésolutionnelle. La TO proposée par Mallat utilise les échelles dyadiques, i.e. qui change pour des résolutions 2^i avec i = 0, 1, 2, ...

La base de cette analyse multirésolutionnelle est une fonction d'échelle, qui est une fonction spline cubique. Une B-spline de l'ordre (2m+1) est générée par les convolutions répétées (2m+1) fois de la B-spline d'ordre 1, qui est une fonction rectangle. La B-spline est une fonction lissée ayant une forme de cloche quand $n \geq 1$, mais moins compacte avec un support de (2m+1). La B-spline d'ordre 2m+1 a des dérivées continues de l'ordre 1 jusqu'à l'ordre 2m-1. On suppose d'abord que la fonction est centrée à l'origine.

La transformée de Fourier (TF) de la B-spline d'ordre (2m+1) est égale à

$$\hat{\Phi}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2} \right)^{2m+1}$$

Pour chaque résolution donnée, les fonctions d'échelle avec les translations discrètes constituent une série de fonctions formant une base. La relation entre ces fonctions dans les deux résolutions voisines est définie comme la relation entre deux échelles qui est exprimée dans le domaine de Fourier comme:

$$\hat{\Phi}(2\omega) = e^{-jk\omega} P(\omega) \hat{\Phi}(\omega)$$

où exp(jkω) représente une translation dans le domaine de l'espace. La fonction d'ondelette est elle aussi déterminée par la relation entre deux échelles comme:

$$\hat{\Psi}(2\omega) = e^{-jk\omega} Q(\omega) \hat{\Phi}(\omega)$$

Comme la fonction d'échelle Φ(ω) est déjà donnée plus haut, il est facile de vérifier que

$$P(\omega) = e^{j\omega/2} (\cos(\omega/2))^{2m+1}$$

et

$$Q(\omega) = 4j\, e^{j\omega/2} \sin(\omega/2)$$

satisfont les relations entre deux échelles si la fonction d'ondelette est donnée par:

$$\hat{\Psi}(\omega) = i\omega \left(\frac{\sin(\omega/4)}{\omega/4} \right)^{2n+2}$$

On peut démontrer que P(ω) et Q(ω) dans les relations données auparavant sont justement la TF de nos filtres digitaux p(n) et q(n) dans le plan de l'image[33]. Ainsi, nous avons deux filtres p(n) et q(n) avec les coefficients:

$p_{-1} = 0{,}125$

$p_0 = 0{,}375$ $\quad q_0 = -2$

$p_1 = 0{,}375$ $\quad q_1 = 2$

$p_2 = 0{,}125$

En observant les filtres digitaux p(n) et q(n), on constate qu'ils ne sont pas orthogonaux. La reconstitution de l'image à partir de la TO doit alors se faire à partir d'une autre paire de filtres. Les filtres d'analyse ne sont pas nécessairement les mêmes que les filtres de synthèse. Ce principe est à la base des algorithmes de TO biorthogonale.

Il est intéressant de noter que l'ondelette de spline quadratique est la première dérivée dans le plan image d'une fonction définie dans le plan de Fourier par:

$$\hat{\Theta}(\omega) = \left(\frac{\sin(\omega/4)}{\omega/4} \right)^{2n+2}$$

Cette fonction fait un lissage de l'image et est une approximation d'une fonction gaussienne. La Figure 59 montrent l'ondelette de spline quadratique, qui est la première dérivée de fonction de spline cubique Q(ω).

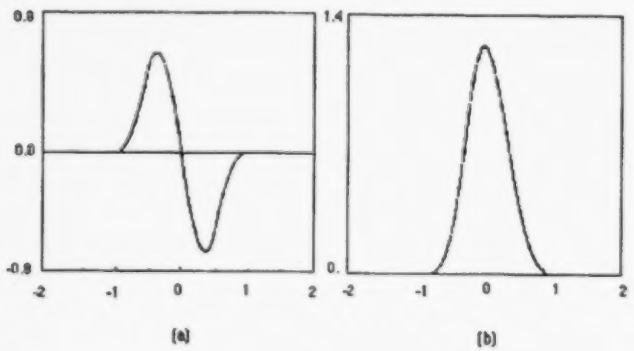


Figure 59. Graphes de l'ondelette de spline quadratique à support compact en (a) et de la spline cubique dont elle est la première dérivée en (b).

Le résultat obtenu par ce filtre, qui par sa nature de première dérivée d'une fonction de lissage donne la meilleure approximation des filtres de contours optimums (dérivée d'une gaussienne), a déjà été présenté auparavant à la Figure 56.

Dans le processus de cascade que nous avons effectué, nous avons dû apporté une normalisation à notre TO pour chaque étape. Ce traitement vise à pouvoir conserver une amplitude uniforme pour les détails importants qui sont retrouvés à plusieurs échelles différentes. Voici les valeurs qui furent utilisées pour l'ondelette de spline quadratique:

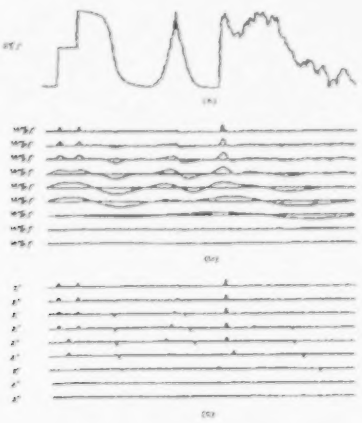
nombre de répétitions:	1	2	3	4	5	6	7
facteur de normalisation:	1,50	1,12	1,0	1,01	1,0	1,0	1,00

11.2.2 Avantages de la détection de bords avec la transformée en ondelette

A première vue, le filtrage passe-haut, q(n) est similaire aux détecteurs de bord de Prewitt, Sobel et Frei-Chen, mais la TO utilise une paire de filtres p(n) et q(n). La TO est une analyse multirésolutionnelle, l'image étant lissée à plusieurs échelles et la détection des bords se faisant sur ces images lissées. De plus, cette détection de bords est moins sensible au bruit et aux petites variations en intensité de l'image. Les filtres p(n) et q(n) sont discrets, mais ce ne sont pas de simples filtres numériques qui calculent les dérivées ou les moyennes de dérivées.

Nous aurions pu avoir recours à la technique du codage sous-bande (subband coding). Cette méthode utilise les filtres numériques sur de multiples bandes de fréquence. Ce codage est d'ailleurs très utile dans la compression de signaux numériques. Cependant, la TO nous offre en plus la condition de régularité. Cette propriété assure que la TO est une opération locale autant dans le plan spatial que dans le plan des fréquences. Ainsi, cette nouvelle théorie mathématique nous permet de procéder à une analyse des fonctions mathématiques, autant dans le cas continu que dans le cas discret.

Mallat a démontré qu'on quantifie la singularité d'une fonction par sa mesure de régularité de Lipschitz a, où a est l'ordre exponentiel qui borne supérieurement la déviation de la fonction. Lorsque $0 < a < 1$, la singularité décrit une fonction de bords tandis que le bruit aléatoire a une mesure de Lipschitz $a < -1$. Comme conséquence, les coefficients de la TO restent approximativement constants pour un grand nombre d'échelles de l'ondelette aux alentours des bords tandis que les coefficients de la TO du bruit vont diminuer avec l'échelle de l'ondelette, comme montré dans la Figure 60. Le contour de l'image et le bruit occupent différentes bandes de fréquence. La TO multirésolutionnelle est capable de les séparer soit par une sélection de bandes de fréquence, soit par une fusion d'informations provenant de plusieurs échelles de la TO.



nombre de répétitions	1	2	3	4	5	6	7
facteur de normalisation	1,50	1,12	1,0	1,01	1,0	1,0	1,00

11.2.2 Avantages de la détection de bords avec la transformée en ondelette

A première vue, le filtrage passe-haut, q(n) est similaire aux détecteurs de bord de Prewitt, Sobel et Frei-Chen, mais la TO utilise une paire de filtres p(n) et q(n). La TO est une analyse multirésolutionnelle, l'image étant lissée à plusieurs échelles et la détection des bords se faisant sur ces images lissées. De plus, cette détection de bords est moins sensible au bruit et aux petites variations en intensité de l'image. Les filtres p(n) et q(n) sont discrets, mais ce ne sont pas de simples filtres numériques qui calculent les dérivées ou les moyennes de dérivées.

Nous aurions pu avoir recours à la technique du codage sous-bande (subband coding). Cette méthode utilise les filtres numériques sur de multiples bandes de fréquence. Ce codage est d'ailleurs très utile dans la compression de signaux numériques. Cependant, la TO nous offre en plus la condition de régularité. Cette propriété assure que la TO est une opération locale autant dans le plan spatial que dans le plan des fréquences. Ainsi, cette nouvelle théorie mathématique nous permet de procéder à une analyse des fonctions mathématiques, autant dans le cas continu que dans le cas discret.

Mallat a démontré qu'on quantifie la singularité d'une fonction par sa mesure de régularité de Lipschitz a, où a est l'ordre exponentiel qui borne supérieurement la déviation de la fonction. Lorsque $0< a < 1$, la singularité décrit une fonction de bords tandis que le bruit aléatoire a une mesure de Lipschitz $a < -1$. Comme conséquence, les coefficients de la TO restent approximativement constants pour un grand nombre d'échelles de l'ondelette aux alentours des bords tandis que les coefficients de la TO du bruit vont diminuer avec l'échelle de l'ondelette, comme montré dans la Figure 60. Le contour de l'image et le bruit occupent différentes bandes de fréquence. La TO multirésolutionnelle est capable de les séparer soit par une sélection de bandes de fréquence, soit par une fusion d'informations provenant de plusieurs échelles de la TO.

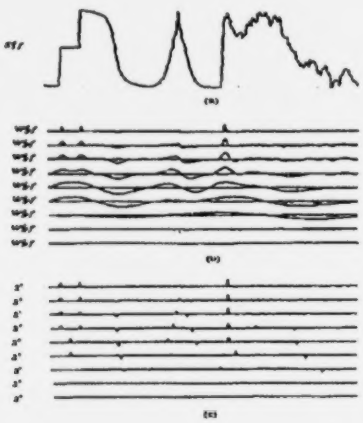


Figure 60. En (a), représentation d'un signal unidimensionnel, en (b) TO discrète du signal selon des échelles dyadiques, en (c) maxima des modules de la TO représentée en (b).

Mallat a démontré que l'image originale peut être reconstituée parfaitement à partir des contours extraits par la TO multirésolutionnelle en utilisant une paire de filtres différente de p(n) et q(n). L'information ne subit pas de pertes suite à l'extraction des contours avec la TO.

11.2.3 Maxima locaux dans la direction du gradient.

Dans le cas d'une image bidimensionnelle, il arrive souvent que les discontinuités ne soient pas isolées, mais qu'elles appartiennent à certains contours de l'image. Le long de ces contours, l'intensité de l'image est discontinue dans une direction, mais varie de façon continue dans la direction perpendiculaire. Ces comportements constituent une information importante de la structure de l'image. A partir de l'information recueillie par la TO, il nous fallait donc déterminer les maxima des coefficients de TO, et ensuite organiser les maxima locaux pour découvrir les courbes de contours.

Pour ce faire, nous avons utilisé les deux composantes de la transformée en ondelette (verticales et horizontales). Nous considérons ces dernières comme un vecteur de gradient[37]. Nous avons défini deux nouvelles matrices. L'une donne le module de ce vecteur gradient, i.e. chaque pixel a une valeur correspondant à:

$$M(i,j) = \sqrt{W_x^2(i,j) + W_y^2(i,j)}$$

L'autre matrice porte, à chacune de ses entrées, l'angle du vecteur gradient calculé comme:

$$A(i,j) = \arctan\left[\frac{W_y(i,j)}{W_x(i,j)}\right]$$

où W_x (i,j) est le coefficients de la TO horizontale du pixel à la position (i,j) et W_y (i,j) dénote la valeur de la TO verticale.

Munis de ces deux matrices d'information, nous avons pu procéder à l'examen qui nous a permis de ne conserver que les maxima locaux dans la direction du gradient. En effet, à chaque pixels (i,j) nous avons quantifié la fonction d'angle A_{ij} en 4 directions (l'axe horizontal, l'axe vertical, l'axe de pente 1 et l'axe de pente -1) et nous avons conservé un pixel (i,j) uniquement si M_{ij} était de valeur supérieure aux pixels voisins, dans une fenêtre 3x3, dans l'axe associé à A_{ij}.

Ce traitement effectué, le résultat obtenu est une image de contours d'épaisseur minimale, soit un pixel, montrée dans Figure 61. Cette condition nous est dictée par le traitement que nous ferons de notre contour par les descripteurs de Fourier. Il est intéressant de noter que d'autres opérateurs, comme la segmentation BCS utilisée par le groupe de Hughes Aircraft, donne des contours d'une épaisseur moyenne de 2,5 pixels. On voit alors que cette technique par la TO est mieux adaptée à notre problème. Lorsqu'appliquée à une image propre, sans background ni bruit aucun, cette technique nous donne d'ailleurs un contour complet de l'objet ciblé. Cependant, avec les images qu'il nous faut traiter, nous obtenons plutôt un contour bruité et non-fermé, qui nécessitera d'autre traitements afin d'obtenir un résultat pouvant être caractérisé par les descripteurs de Fourier.

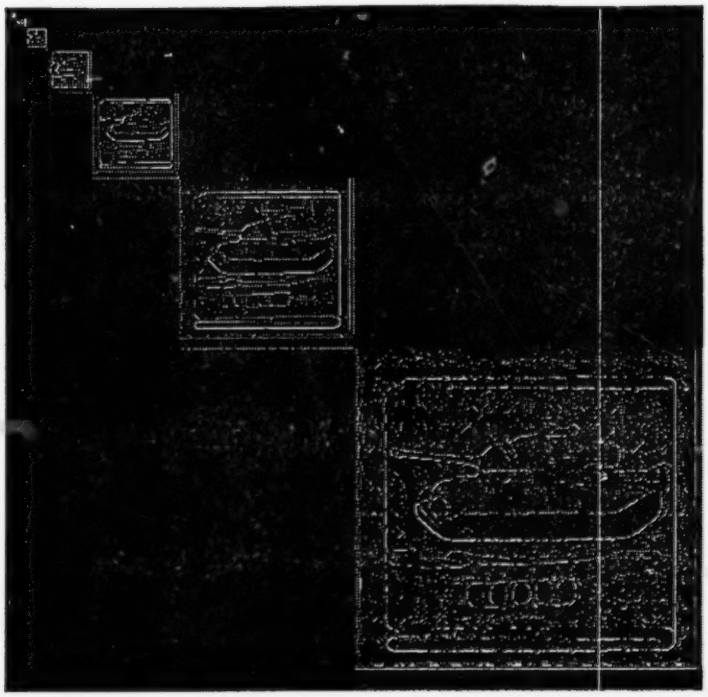


Figure 61. Maxima locaux dans la direction du gradient obtenus à partir de la TO de spline quadratique.

11.3. Recherche du contour de l'objet à analyser.

11.3.1 Segmentation du plateau

À partir de l'image des maxima locaux obtenue, nous avons eu à choisir la résolution que nous allions utiliser pour calculer éventuellement les descripteurs de Fourier. Puisque nous voulions des contours représentatifs de nos objets sans pour autant retenir les plus infimes détails du contours, nous avons choisi la résolution 128x128. Cette réduction de dimensionalité nous a permis également d'accélérer le temps de calcul de notre algorithme. À l'intérieur de cette résolution, nous avons également procédé à une segmentation nous permettant d'isoler les véhicules du plateau rotatif sur lequel ils étaient stationnés ainsi que les tables d'échelle thermique. Cette segmentation ne se fit pas sans heurt car le plateau était en forme de soucoupe, nous obligeant ainsi à conserver des détails nuisibles lors de l'extraction d'une partie rectangulaire de l'image de départ.

11.3.2 Seuillage

Ensuite, en nous basant sur le principe que le maximum local le plus élevé, i.e. la variation d'intensité lumineuse de notre image, doive faire partie du contour recherché, nous l'avons choisi pour être le point de départ de notre tracé de contour. De plus, nous avons procédé à un seuillage nécéssaire pour éliminer les maxima de faible intensité. Bien que l'image soit déjà lissée par les filtres passe-bas p(n), il reste encore des discontinuités à l'intérieur de l'objet et dans la scène de

fond. Ce seuillage rend notre algorithme sensible à la variation des contrastes locaux de l'image, comme cela a été discuté dans le cas du détecteur de Canny. Le seuillage que nous avons considéré, pour être robuste aux variations d'intensités apportés par l'utilisation du détecteur, a été choisi comme étant la moyenne des pixels de l'image des maxima locaux. Ce choix s'est parfois avéré critique dans certains cas lors de l'application de notre algorithme. Nous en reparlerons plus en détail plus tard.

11.3.3 Tracé de courbes

Ayant en main nos images segmentées contenant les pixels de contour d'intensités supérieures à un certain seuil, en plus de connaître notre point de départ, nous étions prêt à effectuer le tracé de la courbe. Dans les techniques de tracé de courbe, il existe deux outils pouvant nous servir: les intensités des pixels et les angles de gradient. Nous utiliserons donc la matrice A vue précédemment ainsi qu'une nouvelle matrice M' ne contenant que les maxima locaux que nous avons conservés.

L'algorithme que nous avons développé suit le contours de la façon suivante. Partant du pixel de départ préalablement trouvé, nous avons tout d'abord examiné la direction α qui est perpendiculaire à l'angle de gradient à ce point. Puisque le point suivant ne peut être qu'un des pixels adjacents, nous avons divisé l'intervalle entre 0 et 2π en huit sous-intervalles. Nos images renfermant un grand nombre de lignes horizontales et verticales, les intervalles associées aux directions gauche, droite, haut et bas seront plus grands que ceux alloués aux directions diagonales, tel que montrer dans la Figure 62.

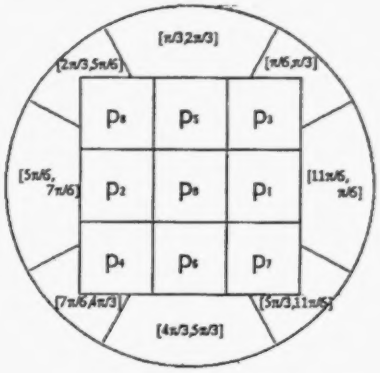


Figure 62. Fenêtre 3x3 montrant la discrétisation de l'angle de direction du tracé et les pixels correspondants.

Nous regardons donc le pixel voisin qui se trouve dans la direction α. Si ce pixel est supérieur au seuil fixé, alors il est considéré comme faisant partie du contour et devient notre nouveau point d'intérêt. Toutefois, si la condition de seuillage n'est pas rencontrée, nous examinons les 2 pixels dans les directions adjacentes au sous-intervalle comprenant α. Ainsi, celui de ces deux pixels ayant la plus grande intensité deviendra notre point de contour. A titre d'exemple, si nous considérons un α compris dans l'intervalle $[11\pi/6,\pi/6]$ et que le pixel p_1 est inférieur au seuil, nous chercherons la suite du contours parmi les points p_3 et p_7. Si ces deux autres pixels ne répondent pas à notre condition, alors on continue à investiguer les intervalles suivants deux par deux, en excluant cependant la possibilité que le pixel soit situé dans la plage située à 180 degrés

de la valeur de la direction données par la matrice A. Dans notre exemple, cela signifie que si p_3 et p_7 sont inférieurs au seuil, nous examinons p_5 et p_6, et ainsi de suite, en excluant la possibilité que p_2 soit conservé. De plus, un pixel déjà répertorié ne peut être choisi comme point de contour, la raison étant que le traitement par les descripteurs de Fourier demande une courbe simple et fermée. Le procédé se poursuit jusqu'à ce que la courbe rejoigne le pixel de départ ou qu'aucun pixel ne satisfasse aux conditions imposées. Dans ce dernier cas, nous avons inséré une sous-routine qui permet d'investiguer le voisinage du pixel du bout pour trouver un raccordement convenable. La recherche se fait successivement à l'intérieur de fenêtres 5x5, 7x7 et 9x9 en arrêtant dès qu'une réponse positive est donnée.

11.3.4 Problèmes rencontrés

Malgré tous les ajustements faits à cet algorithme, nous n'avons malheureusement pas réussi à obtenir, de façon réguilière, des contours simples fermés. Souvent le tracé prenait une direction erronée. Ce phénomène est dû en grande partie au fait que les images ont des contrastes locaux très difficiles à percevoir, même par un observateur humain. Souvent, les mauvais chemins empruntaient les contours du plateau rotatif qu'il ne nous fut pas possible d'éliminer complètement. De plus, le très faible contraste de certaine caractéristique importante du véhicule nous a privé d'importantes informations après avoir été éliminée lors du seuillage. C'est notamment le cas pour le canon qui a une faible intensité dans l'image infrarouge puisqu'il est à la température ambiante. Nous avons donc dû effectuer certaines manipulations par un opérateur humain extérieur, ce qui fait que la méthode n'est pas entièrement automatique. Ces méthodes consistent en trois interventions. D'abord, nous avons quelques fois changé le seuillage pour ne pas perdre d'information. Ensuite, il nous a fallu parfois rehausser la valeur de certains pixels qui permettaient de suivre efficacement le contour voulu. Finalement, nous avons parfois éliminé certains pixels décrivant le plateau ou quelqu'autre information n'ayant pas pu échapper à la segmentation.

Comme nous en avons discuté auparavant, ces interventions d'un opérateur extérieur pourrait être évitées si on ajoute un prétraitement de l'image afin de rehausser ses contrastes locaux. L'utilisation d'un seuillage hystérésis proposé par Canny pourrait aussi aider à tracer les contours. La fermeture du contour pourrait être entièrement automatique. Dans le pire des cas, un algorithme pourrait tracer une ligne droite reliant les deux bouts d'une courbe ouverte. Cependant, dans ce cas, il y a une perte importante d'information.

Après ce travail, nous avons donc réussi à compiler les contours simples fermés pour les cinq appareils ayant une banque complète d'images à tous les 5 degrés, soit le BTR-152, le CHIEF, le GRIZZLY, le LEOPARD et le PT-76. Bien que nous ayons également trouvé les contours fermés pour le AMX, nous ne l'utiliserons pas lors de la reconnaissance, n'ayant que le quart des soixante-douze vues de l'appareil. La Figure 63 montre quelques contours que nous avons obtenus pour les cinq véhicules.

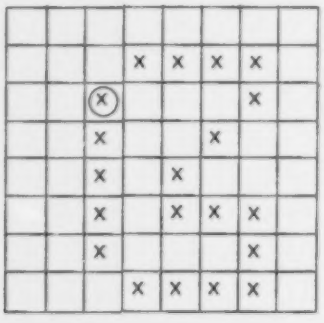
Figure 63. Contours de quelques vues des cinq véhicules qui ont servis à notre étude.

12. Description du contours

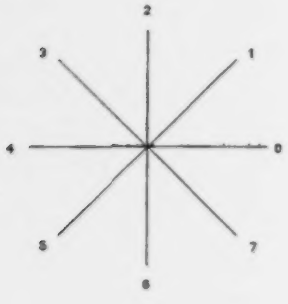
Une fois que le contour est extrait de l'image, la prochaine étape est sa description par un nombre limité de paramètres numériques. Ainsi, la reconnaissance des objets 3D pourra être réalisée par une classification dans un espace de caractéristiques. Les descripteurs de l'image donnent aussi certaines invariances sous les transformations de similarité (translation, rotation dans le plan et changement d'échelle). Plusieurs méthodes de description du contour existent, parmi lesquelles le codage en chaîne, le descripteur de Fourier-Mellin et le descripteur de Fourier sont généralement utilisés.

12.1. Codage en chaîne

Le codage en chaîne est une représentation de forme basée sur le contour. La Figure 64 montre le principe de ce codage.



(a)



(b)

1 0 0 0 6 5 5 6 0 0 6 6 4 4 4 3 2 2 2 2

(c)

Figure 64. Représentation du codage en chaîne d'une courbe digitale. (a) Courbe digitale (le X encerclé est le point de départ). (b) Directions du codage en chaîne. (c) Code pour la courbe en (a).

Les codes sont les pentes intrinsèques de la courbe. Il y a seulement huit pentes possibles pour une paire de points sur la courbe. Ces huit pentes sont représentées par huit chiffres. Ainsi, une courbe peut être représentée comme une chaîne de codes. L'avantage du codage en chaîne est la simplicité de la méthode. Un contour 2D est décrit par une série de nombres entiers, un par pixel. Cette compression de données favorise les traitements ultérieurs car celles-ci prennent moins d'espace dans la mémoire d'un ordinateur. Ce fait est important dans le cas de la reconnaissance d'objet 3D.

Les codes en chaîne sont invariants sous translation, mais ne sont pas invariants sous rotation et changement d'échelle. Sadjadi a proposé l'utilisation de l'histogramme des codes en chaîne. Un changement d'échelle se traduit par une translation dans le sens vertical de l'histogramme et une rotation se traduit par une translation cyclique horizontale dans l'histogramme. Avec cette technique, Sadjadi a réalisé la reconnaissance de cibles tactiques dans des images réelles. La classification est robuste pour les cibles ayant jusqu'à 10% de sa superficie camouflée par des obstacles[38]. Toutefois, cette approche a un désavantage. En effet, la description d'une courbe par l'histogramme des codes de chaîne est moins précise que celle par les codes de chaîne eux-même. On perd une partie de la capacité de discrimination en passant de l'une à l'autre.

12.2. Descripteur de Fourier-Mellin

Le descripteur de Fourier-Mellin (DFM) est une représentation de forme basée sur toute la région décrivant l'objet (contour avec intérieur rempli) . Il utilise les moments de l'image définis dans un système de coordonnées polaires. Le DFM est formé de paramètres numériques invariants sous rotation dans le plan de l'image et changement d'échelle[39]. Il est moins sensible au bruit de fond que les invariants de moments proposés par Hu[40]. Le DFM orthogonal permet de choisir les ordres des moments dans un polynôme radial pour une meilleure représentation de l'image avec moins de paramètres[41].

Le DFM est appliqué à une image segmentée. Dans le cas de notre projet, les images infrarouges de véhicules doivent être binarisées. Cette étape doit être effectuée pour avoir la même intensité à l'intérieur de l'image, car des changements de température pour différentes parties du véhicule peuvent changer la luminance de l'image et donc changer son DFM.

La méthode des moments est couramment utilisée dans la description d'image et la reconnaissance des formes. Elle pourra être utilisée pour décrire les images après la segmentation, le rehaussement du contraste et la binarisation. Pour une image de basse résolution, le contour de l'image peut être très accidenté. Le DFM conviendra alors mieux qu'un descripteur du contour, comme le descripteur de Fourier.

12.3. Descripteur de Fourier (DF)

Le DF fournit une méthode populaire pour représenter le contour d'une forme 2D. Une courbe planaire fermée peut être représentée par une fonction complexe périodique $z(\gamma)$ d'un paramètre continu: la longueur d'arc γ de la courbe.

$$z(\gamma) = x(\gamma) + jy(\gamma)$$

ou par un angle cumulatif $\phi(\gamma)$, définit par

$$\phi(\gamma) = \theta(\gamma) - \theta(0)$$

qui est la mesure nette de l'angle entre le point de départ $\gamma = 0$ et le point dont la longueur d'arc est égale à γ ($\theta(\gamma)$ est l'angle tangent à la courbe au point γ). Le point de départ est arbitrairement choisi. Soit L la longueur d'arc totale du contour; on voit facilement que $\phi(0) = 0$ et $\phi(L) = -2\pi$. La fonction $\phi(\gamma)$ est périodique de période de L.

On définit un paramètre t qui varie de 0 à 2π et on définit la fonction suivante:

$$\phi^*(t) = \phi(\frac{L}{2\pi}t) + t$$

On voit aisément que $\phi^*(0) = \phi^*(2\pi) = 0$. Toutes les courbes ayant une forme identique et le même point de départ ont la même fonction $\phi^*(t)$. Les invariances sous transformations de similarité (translation, rotation dans le plan et changement d'échelle) sont automatiques, vu l'utilisation des angle tangents à la courbe. On voit directement que la fonction $\phi^*(t)$ est invariante sous translation. Elle est aussi invariante sous changement d'échelle, c'est-à-dire, sous changement de L car la courbe est une fonction périodique de γ avec une période L et elle est ensuite reparamétrisée sur $[0,2\pi]$.

Pour enlever l'influence du choix de point de départ, on utilise la série de Fourier de $\phi^*(t)$ qui est une fonction périodique de période 2π:

$$\phi^*(t) = \mu_0 + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \sin kt)$$

$$= \mu_0 + \sum_{k=1}^{\infty} A_k \cos(kt - \alpha_k)$$

Le descripteur de Fourier[42] est défini par les amplitudes des coefficients de Fourier A_k et les phases α_k. Si l'image tourne mais le point départ ne la suit pas, la fonction $\phi^*(t)$ va être changée par une constante. Ses coefficients de Fourier vont être changés par une phase constante, mais les amplitudes A_k seront invariantes sous rotation, translation, changement d'échelle et choix du point de départ.

La méthode d'angle cumulatif proposée par Zahn et Roskies a pour avantage d'être moins sensible aux variations des hautes fréquences de l'angle tangent à la courbe. Celles-ci correspondent aux détails de petite dimension de la courbe et ne changent pas l'allure du contour. Ceci est très intéressant pour notre projet car les petits détails des contours sont associés au bruit de détection thermique qui n'apportent pas d'information utile pour la classification des véhicules. Le principal désavantage de la méthode d'angle cumulatif est qu'elle demande le calcul des dérivées de la courbe.

Le DF représente une compression de données importante. Il contient un nombre restreint de paramètres très condensés qui décrivent l'objet. La dimension du vecteur de DF est très petite par rapport à la dimension de l'image 512 x 512. Les DFs sont fondés sur le contour de l'image, qui porte l'information essentielle de l'objet. D'après notre expérience, le DF est beaucoup moins sensible que d'autres méthodes, comme le filtre d'harmonique circulaire et le descripteur de Fourier-Mellin, aux déformations de l'image causées par une rotation hors-plan de l'objet. On pourra donc décrire un objet tourné hors-plan par ses images dans un intervalle de 5 à 15 degrés de rotation. Les DFs permettent non seulement de diminuer la dimension mais aussi le nombre de vecteurs qui vont être stockés dans le réseau de neurones qui servira à la reconnaissance des formes d'un objet tridimensionnel.

12.3.1 Evaluation des descripteurs de Fourier

12.3.1.1 Elimination des points isolés et redondants

Le calcul des DFs requiert que l'image soit représentée par son contour externe. Ce contour doit être une courbe simple fermée d'un pixel d'épaisseur. Nous avons donc dû procéder à un traitement de notre courbe visant à réduire le nombre de pixels de contours à 3 dans une fenêtre 3x3. Grâce à ce traitement, nous pourrons calculer les tangentes en prenant l'angle formé par le segment passant par les deux points voisins du pixel étudié. En faisant l'examen de chaque point faisant partie du contour, nous avons posé nul tout pixel n'ayant qu'un seul voisin d'intensité non-nulle dans la fenêtre l'entourant. En effet, il correspond alors à une portion de contour inutile décrivant une déviation temporaire de notre parcours fermé en dehors de notre courbe d'intérêt. Un pixel n'ayant qu'un seul voisin pourrait être considéré comme un bout de courbe ouverte et nous devons avoir une courbe fermée pour procéder avec la méthode.

Ensuite, nous avons analysé le cas très fréquent où il y a plus de trois pixels dans la fenêtre entourant le pixel étudié. Ce cas rend problématique l'évaluation de la tangente en un point. Pour réussir à connaître les deux pixels adjacents à conserver, nous avons élargi notre étude à la fenêtre 5x5 entourant notre cellule traitée. L'algorithme que nous avons construit donne la préséance aux pixels disposés sur les directions diagonales. Cette décision résulte d'une étude systématique des différents cas où il y a plus de 3 pixels dans la fenêtre 3x3. Ce traitement fait, nous sommes alors prêt à procéder au calcul des descripteurs de Fourier.

12.3.1.2 Calcul des descripteurs de Fourier

Muni de tous les contours des 72 vues des 5 chars étudiés, il nous était possible d'évaluer les descripteurs de Fourier. La fonction d'angle tangent fut calculée à chaque pixel, cette valeur étant trouvée grâce à la position des deux pixels voisins. Lors de l'évaluation de la fonction à un pixel donné, la longueur relative de la courbe fut augmentée de 1 si le pixel précédant se trouvait dans une direction

verticale ou horizontale alors qu'un accroissement de $\sqrt{2}$ 0 fut apporté lorsqu'il se trouvait dans une direction diagonale.

La fonction $\phi^*(t)$ est ainsi échantillonnée sur des intervalles de longueur d'arc non-uniformes. L'utilisation de la transformée de Fourier rapide (FFT) n'est donc pas possible. Nous avons alors tout simplement utilisé la forme trouvée précédemment pour la TF de la fonction d'angle $\phi^*(t)$. Les coefficients a_k 0et b_k 0sont alors données par les relations:

$$a_k = \frac{1}{\pi}\int_0^{2\pi} \phi^*(t)\cos kt dt$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} \phi^*(t)\sin kt dt$$

Si nous exprimons ces intégrales en fonction de l'angle cumulatif ϕ, nous avons la relation:

$$a_k = \frac{2}{L}\int_0^{L} (\Phi(l) + 2\pi l / L)\cos(2\pi kl / L) dl$$

$$b_k = \frac{2}{L}\int_0^{L} (\Phi(l) + 2\pi l / L)\sin(2\pi kl / L) dl$$

Finalement, en se rappelant que la fonction d'angle cumulatif est une fonction discrète à échelons ($\Phi 0 = \Phi_n$ avec Φ_n 0 constant sur l'intervalle $[l_n, l_{n+1}]$, où l_n 0donne la longueur cumulative de la courbe entre le point de départ et le point n0) et en effectuant le changement de variable $\xi = 2\pi\ kl/L$, on peut écrire:

$$a_k = \frac{2}{L}\frac{L}{2\pi k}\left[\sum_{n=1}^{m}\int_{\xi_{n-1}}^{\xi_n} \phi_{n-1}\cos(\xi)d\xi + \frac{1}{k}\int_0^{2\pi k} \xi\cos(\xi)d\xi\right]$$

$$b_k = \frac{2}{L}\frac{L}{2\pi k}\left[\sum_{n=1}^{m}\int_{\xi_{n-1}}^{\xi_n} \phi_{n-1}\sin(\xi)d\xi + \frac{1}{k}\int_0^{2\pi k} \xi\sin(\xi)d\xi\right]$$

où on a défini $\xi_n = 2\pi\ nl_n/L$. L'intégrale de $\xi\cos(\xi)$ est nulle et celle de $\xi\sin(\xi)$ est égale à -2π et peut, suite à de longs développements mathématiques[42], être additionnée au terme constant μ_0 de l'expansion de Fourier. En sortant les ϕ_{n-1} des premières parties des deux équations, on intègre et, en remaniant les termes, on trouve la forme discrète pour les coefficients:

$$a_k = \frac{-1}{k\pi}\sum_{n=1}^{m}\Delta\Phi_n \sin\left[\frac{2\pi k}{L} l_n\right]$$

$$b_k = \frac{1}{k\pi}\sum_{n=1}^{m}\Delta\Phi_n \cos\left[\frac{2\pi k}{L} l_n\right]$$

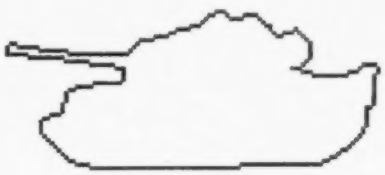
où $\Delta\phi_n$ est le changement d'angle tangent entre le pixel étudié et le précédent sur la courbe. On calcule les amplitudes facilement en calculant $A_k=[a_x^2+b_k^2]^{1/2}0$.

Nous n'avons conservé que les amplitudes de la transformée de Fourier car ce sont elles qui sont invariantes sous choix du point de départ, translation, échelle et rotation de la courbe dans le plan. Les phases, pour avoir de telles caractéristiques, doivent être remaniées sous la forme de différences entre chaques composantes retenues dans l'expansion de Fourier comme suit:

$$F_{kj}[\gamma]=j^*\alpha_k-k^*\alpha_j$$

où $j^*=j/\mathrm{gcd}(j,k)$ et $k^*=k/\mathrm{gcd}(j,k)$ et gcd(j,k) représente le plus grand commun diviseur de j et k.

On voit donc que pour avoir les mêmes invariances que le vecteur d'amplitudes, on doit définir une matrice Fkj, élevant au carré le nombre d'éléments nécéssaires pour décrire notre contour. Par exemple, pour un vecteur de 25 amplitudes, on obtient une matrice de phases de 25x25=625 éléments. Une autre raison qui nous a incité à ne pas considérer cette information d'angle est qu'elle sert principalement à reconnaître deux images miroirs l'une de l'autre, ce qui n'est pas pertinent pour les objectifs visés par notre projet.



A_1= 0,177750	A_{40}= 0,006994
A_{10}= 0,093733	A_{50}= 0,029285
A_{20}= 0,063944	A_{60}= 0,020080
A_{30}= 0,033189	A_{70}= 0,014053

Figure 65. Contour d'un char CHIEF à un angle de 100 degrés et quelques uns de ses DFs associés.

12.3.1.3 Normalisation des descripteurs de Fourier

Nous avons pu remarquer que l'ordre de grandeur des amplitudes A_k0 varie selon k . Ce fait est bien montré dans la Figure 65 nous montrant un contour de char avec quelques uns de ses DFs. Ainsi, dans certains processus de reconnaissance, tels ceux basés sur la distance euclidienne minimale entre deux vecteurs, les composantes de différents k auront des contributions différentes. Il est possible que ce comportement soit très bien adapté à ce type de caractérisation de nos contours, mais il pourrait éventuellement être souhaitable que chaque composante du vecteur soit d'un ordre de grandeur similaire. Pour ce faire, nous avons compilé une deuxième série de descripteurs de Fourier normalisés décrivant nos objets. Comme facteur de normalisation pour une amplitude, nous avons compilé la variance de la

liste de toutes les entrées A_k de notre banque de descripteurs de Fourier, i.e. les 72 projections des 5 véhicules étudiés. Pour l'amplitude A_k d'un véhicule en particulier, nous avons donc calculé cette variance σ_k sur les valeurs A_k des descripteurs de la banque de donnée. Les nouveaux DFs ainsi obtenus nous serviront à définir un deuxième espace caractéristique afin de procéder à la reconnaissance.

Dans le processus de reconnaissance qui suivra, nous verrons qu'une fois les descripteurs calculés, il nous reste un paramètre libre à ajuster pour obtenir la reconnaissance optimale, soit le nombre de descripteurs formant les vecteurs caractéristiques de notre espace de reconnaissance.

13. Reconnaissance d'objet 3-D

Notre tâche consiste à reconnaitre des objets 3-D à partir de leurs images bidimensionnelles. Nous ne disposons pas d'information sur la troisième dimension, i.e. la profondeur de l'objet. L'objet 3-D est donc représenté par un très grand nombre d'images 2-D avec des changements d'échelle et des rotations dans le plan et en dehors du plan de l'image. On veut caractériser l'information 3-D à partir de ses images 2-D, mais une partie importante de l'information concernant l'objet est perdue par ce passage à une dimension inférieure.

La reconnaissance de l'objet tridimensionnel à partir de son image de projection bidimensionnelle est un problème contenant des failles (ill-posed problem). Un problème est bien défini si il a une solution et que cette dernière est unique. De plus, cette solution dépend des données initiales de façon continue. Nous disons donc que notre problème contient des failles parce qu'un nombre infini de solutions d'objet 3-D existent pour la même image de projection.

On peut voir intuitivement la difficulté de notre tâche. Quand l'objet tourne en dehors du plan, la variation de ses images de projection est très importante. La différence entre les images de projection du même objet est souvent plus importante que la différence entre les images de différents objets vus du même point d'observation. L'autre difficulté vient du fait que la reconnaissance de l'objet 3-D doit être invariante sous translation, changement d'échelle et rotation dans le plan-image (invariance de similarité) et en dehors du plan-image (invariance de point de vue). Pour résoudre ce problème, plusieurs méthodes ont été proposées. En voici quelques unes qui s'ajoutent à celle décrites dans la partie de revue des méthodes existantes.

13.1. Méthode de dessin de lignes

Une fois qu'une image de l'objet est segmentée et que les lignes de bords sont extraites, nous obtenons un dessin de lignes pour représenter cet objet 3-D. Un dessin de lignes est simplement une collection graphique de lignes sur papier, donc en 2-D. C'est une représentation de l'objet qui paraît naturelle pour l'être humain dans la vie de tous les jours. Ces lignes 2-D possèdent une perspective, donnant l'illusion que l'objet est tridimensionnel.

Une approche populaire de la reconnaissance d'objets 3-D consiste à interpréter de multiples dessins de lignes de différentes vues perspectives et d'établir des relations entre les images ou plutôt des correspondances entre les points sur les images. Finalement, on procède à la construction de l'objet 3-D à partir des correspondances de points[43]. Une autre approche consiste à considérer plutôt une seule image et d'imposer ensuite suffisamment d'hypothèses et de contraintes sur l'objet 3-D de telle sorte que le passage de l'image 2-D à l'objet 3-D devient inversible.

Dans notre projet, le profil des véhicules est constitué par des surfaces solides, dont une partie est planaire. Les images sont plutôt des projections orthogonales que des projections en perspectives. Dans le dessin de lignes de l'objet 3-D, les lignes 2-D sont les contours et les discontinuités des surfaces de l'objet. Malheureusement, dans le cas des images infrarouges, l'intensité de l'image dépend davantage de la température que de l'éclairage. Les lignes d'intersection des surfaces ne paraissent plus dans les images infrarouges; les objets sont opaques. Seules les lignes de contours restent. Alors, la méthode de dessin de lignes devient inefficace à cause de la perte des lignes d'intersection des surfaces.

13.2. Descripteur de Fourier invariant sous transformation affine

Arbter et al. ont proposé un descripteur de Fourier invariant sous les transformations affines[44]. Une transformation affine inclut les transformations de similarité (rotations dans le plan de l'image, changements d'échelle et translations), en plus de la transformation de ciseau. Cette dernière est une distorsion de l'image qui correspond à ce qui se passe quand un objet planaire est observé par une caméra sous des angles d'azimuts arbitraires par rapport au plan de l'objet, comme montré à la Figure 66. La transformation de ciseau est une approximation de la transformation de perspective. Dans ce DF, le paramètre de la courbe γ, qui est la longueur d'arc de la courbe, est remplacé par un paramètre de "longueur affine". Cette technique est utilisée pour la reconnaissance d'avions, qui sont des silhouettes plutôt planaires. Dans notre application les silhouettes des véhicules peuvent être considérées comme planaires seulement quand l'angle azimutal atteint 90 degrés.

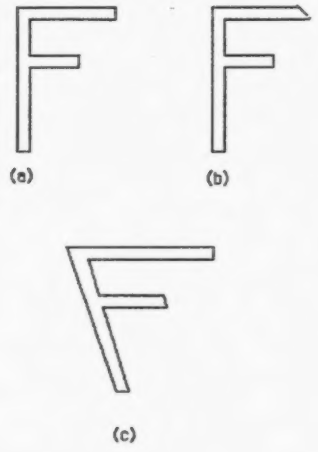


Figure 66. Exemples de transformations. En (a), on voit une lettre F de référence. En (b), la transformation n'est pas affine et en (c), elle est affine et représente le F de (a) comme s'il était vu avec un angle azimuthal.

13.3. Réseaux de neurones

Dans notre application, la reconnaissance de l'objet 3-D est modélisée de la façon suivante: l'objet 3-D est considéré comme une fonction originale n'ayant pas de représentation analytique. Une série d'images 2-D représentent les projections de l'objet à de multiples vues. Ces images sont considérées comme les échantillons de l'objet volumique 3-D.

Nous ne disposons pas de relations géométriques entre les images de projection ni de transformations mathématiques qui permette de passer d'une image à l'autre. L'objet est donc représenté par un très grand nombre d'images 2-D. Il y a deux cheminements pour solutionner notre problème. Dans un premier, on prend toutes les images dans la mémoire. La reconnaissance est effectuée par une correspondance entre l'image d'entrée et les images stockées (template matching), via une mémoire associative ou des filtres composites optimisés. Dans la deuxième approche, on décrit les images par des paramètres numériques en introduisant certaines

invariances. Les objets à reconnaître seront par la suite classés dans l'espace des caractéristiques à l'aide d'un réseau de neurones.

Dans les deux approches, les difficultés demeurent les mêmes. On a un trop grand nombre d'images 2-D représentant un objet lorsque l'objet change d'échelle, d'orientation et de point de vue; on a également de trop grandes variations entre les images. Alors, la réduction du nombre d'images représentatives est importante pour la performance du système de reconnaissance. Les DFs représentent une compression importante des données (de l'image 2-D à une série de 10 à 100 paramètres numériques). De plus, les DFs sont invariants sous les rotations dans le plan de l'image et sous les changements d'échelle, ce qui réduit considérablement le nombre d'images représentant un objet 3-D.

Il existe cinq types de réseaux de neurones qui sont susceptibles d'être utiles pour résoudre notre problème. Un premier est le Néocognitron[45] que l'on a rencontré dans le système SAHTIRN. Un autre est le modèle de description structurale, tel qu'utilisé par le système XTRS du MIT Lincoln Lab. Voyons maintenant un aperçu des trois autres.

13.3.1 Modèle de transformation et matching

Ce modèle nous est familier. Si nous utilisons les coefficients des DFs comme données d'entrée d'un réseau de neurones de type "rétropropagation"[46], le réseau peut effectuer l'apprentissage supervisé pour la reconnaissance de l'objet 3-D. Les DFs sont déjà invariants sous rotations dans le plan de l'image et changements d'échelle. Le réseau doit être capable d'associer à une même classe les images de différents points de vue d'un même objet.

Le problème prévisible dans cette approche est que la différence entre les images de différents points de vue d'un objet peut être plus importante que la différence entre les images de différents objets. Si chaque image est représentée par un point dans l'espace des caractéristiques, le groupe de points d'un objet 3-D a de grandes chances de chevaucher les groupes de points des autres objets. La classification correspond à une division non-linéaire dans l'espace des caractéristiques. Le réseau de neurones de "rétropropagation" a beaucoup de difficulté à former cette division dans ses couches cachées durant l'apprentissage. Ce dernier peut prendre un temps de calcul très long ou ne pas arriver à la classification voulue.

13.3.2 Réseau généralisé de fonctions de base

Ce réseau est proposé par Poggio et al. du MIT, Laboratoire d'intelligence artificielle pour la reconnaissance automatique d'objet 3-D à partir de ses images de projections[47]. Supposons que les objets sont formés des lignes représentant une perspective, un peu comme lorsqu'on dessine un cube en considérant les trois dimensions (type frame-lines). Le théorème d'extraction de la structure d'un objet à partir de ses mouvements[48] indique que l'information complète de l'objet 3-D est présentée par seulement deux vues perspectives, à la condition que l'objet soit représenté par une série de 5 à 8 points caractéristiques, et que ces points soient identifiés dans chacune des vues perspectives. Dans notre cas les objets sont opaques. Les points caractéristiques ne peuvent pas être identifiés dans chaque image infrarouge.

Nous devons reconnaître les objets à partir d'un grand nombre d'images 2-D de projection. La reconnaissance de l'objet 3-D à partir de ces images est un problème contenant des failles car il y a beaucoup d'objets 3-D qui peuvent donner la même image 2-D. L'approche proposé par Poggio

consiste à appliquer certaines contraintes de lissage et à entraîner un réseau de neurones, dont la relation entre l'entrée et la sortie est lissée. Quand on arrive à stabiliser la solution, le problème devient un problème bien défini. L'hypothèse fondamentale de ce modèle de réseau de neurones est que l'image de projection 2-D d'un objet 3-D varie de façon graduelle, continue et lissée lors d'une variation de l'angle de vue.

On veut synthétiser une fonction de plusieurs variables à partir d'un petit nombre d'images de projection. Cette fonction, appliquée à une vue de l'objet, va générer une vue standard de cet objet tandis que son application à une vue d'un différent objet va générer une mauvaise image qui est facilement détectée. Cette fonction est obtenue par l'apprentissage d'un réseau de neurones de fonctions de base radiale.

La première couche du réseau reçoit le vecteur **x** qui représente une vue perspective de l'objet. Le nombre de neurones dans la couche cachée peut être égal au nombre d'images de projection dans le groupe d'apprentissage. Dans la couche cachée, la sortie d'un neurone est donnée par

$$h_j = \Phi_j\left(\left\|\vec{x}-\vec{t}_j\right\|\right)$$

où Φ_j est la fonction de transfert non-linéaire du neurone j, t_j est le prototype de l'image d'apprentissage ou le centre du groupe de vecteurs d'images d'apprentissage. Si Φ_j est une fonction gaussienne avec une non-linéarité localisée et de décroissance exponentielle et $\left\|\vec{x}-\vec{t}_j\right\|$ est la distance euclidienne, h_j est alors une fonction radiale de la distance euclidienne.

La sortie du réseau O_k pour un neurone k de la troisième couche est une combinaison linéaire des fonctions radiales non-linéaires de base.

$$O_k = \sum_j \lambda_{jk} h_j$$

Les poids des synapses λ_{jk} sont ajustables pendant l'apprentissage. Ce dernier est supervisé de façon à ce que la sortie du réseau soit une vue standard de l'objet. Il s'agit en effet d'un processus d'optimisation linéaire. Si $d_{k,i}$ est la valeur de sortie désirée du neurone de sortie k pour l'entrée x_i, l'apprentissage est fait avec une fonction de coût:

$$C(o)=\sum_{i,k}\left(d_{i,k}-\sum_{j=1}^{M}\lambda_j\Phi_j\left\|\vec{x}_i-\vec{t}_j\right\|\right)^2$$

La fonction gausienne Φ est localement centrée sur t_j, qui représente l'une des images d'apprentissage. La fonction de sortie du réseau est une combinaison de gaussiennes centrées sur les images de différentes vues. Quand l'image d'entrée **x** est exactement égale à t_j, nous avons la sortie désirée. Si **x** est différente de t_j, la sortie

$$O_k = \sum_j \lambda_{j,k}\Phi_j\left(\left\|\vec{x}-\vec{t}_j\right\|\right)$$

est une fonction d'approximation de $\|\vec{x}-\vec{t}_j\|$ qui correspond à un lissage par les gaussiennes entre les images d'apprentissage $\mathbf{t_j}$. La sortie sera soit la vue standard désirée, soit quelque chose de complètement différent, dépendamment de la distance $\|\vec{x}-\vec{t}_j\|$.

La fonction de coût définie précédemment est l'erreur classique. Dans les réseaux de neurones régularisés, on ajoute un autre terme dans la fonction de coût, ce qui permettra de stabiliser le lissage de la fonction O. Ce lissage rend le problème bien défini.

Le modèle de Poggio a été démontré avec des images de type "frame-lines". Une image de vue est décrite par de cinq à huit paires de points caractéristiques et le modèle est réalisé avec un petit nombre d'images. Par conséquent, la dimension du réseau est relativement petite. Il serait intéressant d'appliquer le modèle de Poggio à nos problème de reconnaissance à partir des contours d'images.

13.3.3 Modèle de trajectoire dans l'espace des caractéristiques (TEC)

13.3.3.1 Description de la méthode

Le modèle de réseau de neurones de trajectoires dans l'espace des caractéristiques (TEC) a été proposé pour la reconnaissance des objets 3-D par Casasent et Neiberg[27]. Ce modèle, basé sur le principe de distance minimale et sur l'interpolation entre les images de projection, est capable de contourner le problème de définition des surfaces de décision non-linéaires dans l'espace des caractéristiques pour la classification entre les groupes de points qui se sont superposés. La méthode est simple, efficace, et capable de donner des résultats satisfaisants.

Voici les grandes lignes de la méthode. Nous disposons des images de projection des véhicules à classifier. Nous faisons le rehaussement des images, l'extraction des contours et la description des contours avec le DF. Ce vecteur décrit une image de projection dans un espace de caractéristiques de dimension n, où n est le nombre de coefficients composant le DF. A chaque image correspond donc un point dans l'espace des caractéristiques.

Les vecteurs de DF sont déjà invariants sous rotation dans le plan de l'image, translation et sous changement d'échelle. Il nous reste à régler le problème de rotation hors-plan de l'image. Si les images de projection sont prises à de très petits intervalles d'angle de projection, les points caractéristiques des vues perspectives consécutives de l'objet 3-D pourraient être reliés, formant ainsi une courbe dans l'espace des caractéristiques. Si M est le nombre d'objets 3-D à classifier nous aurons M courbes dans l'espace des caractéristiques. Un objet 3-D pourra être classifié avec une de ces images de projection par le principe du plus proche voisin puisque le point caractéristique de cette image se confondrait alors avec l'une des M courbes. Mais, cette approche est irréaliste car elle a besoin d'un très grand nombre d'images de projection dont on ne dispose pas. Le grand nombre de vecteurs nécéssite un réseau de neurones de trop grande dimension et un trop grand espace de mémoire.

Nous voulons former un réseau de neurones avec le moins de points. Les images le composant pourront être, par exemple, les projections des objets à tous les 10 ou 15 degrés. L'hypothèse de base est alors que la variation de l'image de projection est continue et graduelle avec une rotation

hors-plan de l'objet 3-D. Dans le modèle de Casasent, cette propriété est utilisée de façon explicite: on relie simplement les points représentant deux vues consécutives du même objet par une ligne droite dans l'espace des caractéristiques. De plus, sachant que les projections représentent une rotation hors plan de 360 degrés, nous pourrons relier les deux points extrêmes. Un objet 3-D est donc représenté par une courbe continue, fermée, formée de segments de droite et avec des inflexions aux points choisis pour l'apprentissage.

Dans la classification, on calcule les distances entre un point de l'image d'entrée inconnue et tous les segments linéaires des courbes qui représentent les M objets 3-D. La classification est réalisée par le principe de distance minimale. De cette façon, l'espace des caractéristiques est divisé pour la classification en zones qui suivent le trajet de chacune des courbes représentant les objets.

Casasent et al. décrivent cette approche dans un cadre de réseau de neurones, et l'appellent réseau de neurones de trajectoire dans espace des caractéristiques (TEC). La première couche est la couche pour entrer les données. Le nombre de neurones est égale à n, la dimension de l'espace des caractéristiques. La deuxième couche, une couche cachée, représente les paramètres géométriques des segments de droite qui relient les points caractéristiques des vues consécutives du même objet. Le nombre de neurones dans la couche cachée est égale au nombre de segments de droite composant les courbes. Les sorties de la couche cachée sont les distances calculées entre le point d'entrée et tous les segments de droite. La troisième couche est un réseau de gagnant-prend-tout (winner-take-all), qui cherche la distance minimale. L'apprentissage du réseau consiste simplement à emmagasiner les données d'apprentissage provenant de la série d'entraînement des images de projection dans la couche cachée. Dans ce sens on peut dire que ce n'est pas un vrai réseau de neurones. Il s'agit plutôt d'un algorithme de distance minimale avec interpolation linéaire entre les points d'apprentissage.

13.3.3.2 Résultats obtenus

Pour l'apprentissage nous avons considéré les images des projections d'un objet à tous les 15 (ou 10) degrés de rotation hors-plan. Nous avons obtenu les 24 (ou 36) vecteurs de DFs représentant l'un des véhicules. La dimension n du vecteur de DF varie de 10 à 100. Nous relions les points caractéristiques des images de rotation consécutives par des segments de droite. Nous supposons que cette interpolation linéaire décrit bien le comportement des vues intermédiaires de l'appareil. Nous avons testé ceci en calculant le pourcentage de déviation des vues intermédiaires pour un char CHIEF par rapport à la valeur donnée par la courbe. Pour ce faire, nous avons calculé la distance minimale entre le DF des images du CHIEF et la courbe calculée pour ce même véhicule. Ensuite, nous faisons le rapport suivant:

$$D_k = \frac{\min \left\| DF(CHIEF_k) - \rho_{CHIEF} \right\|}{\left\| DF(CHIEF_k) \right\|} \cdot 100\%$$

où k donne la projection que l'on étudie. Nous avons alors un taux de variation nous donnant une appréciation de l'approximation obtenue par l'extrapolation linéaire. Les résultats sont donnés à la Figure 67 pour un apprentissage à tous les 15 degrés avec des DFs de dimension 75.

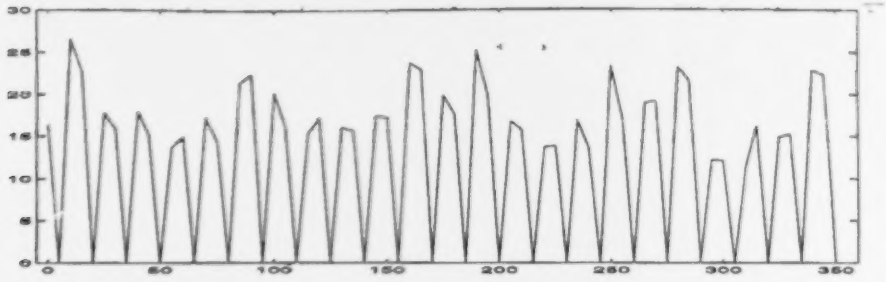


Figure 67. Pourcentage de déviation des vecteurs de DFs (n=75) pour le char CHIEF par rapport à la courbe apprise à tous les 15 degrés.

Nous avons donc obtenu une courbe pour un véhicule qui consiste en 24 (ou 36) segments de droite. Nous travaillons avec cinq véhicules. Ces véhicules sont répertoriés en deux classes: les chars d'assaut avec le CHIEF et le LEOPARD, et les véhicules de transport de troupes, avec le BTR-152, le GRIZZLY et le PT-76. Nous remarquons que le PT-76 est le seul véhicule de transport qui possède un canon. Dans une étape intermédiaire de l'algorithme, nous utilisons 24 x 5 = 120, (ou 36 x 5 = 180) données. Chacun d'elles emmagasine les paramètres d'un segment de droite.

La série d'images servant pour les tests est formée de 72 projections pour un véhicule en rotation hors-plan pour tous les cinq degrés, et nous avons 5 véhicules. Au total, 360 points sont classifiés dans l'espace des caractéristiques. Parmi ces 360 points, il y en a 120 (ou 180) qui appartiennent à la série d'apprentissage. Ils constituent les points d'inflexion des cinq courbes. Le reste des points correspondent aux images de projection intermédiaires n'ayant pas été apprises par l'algorithme.

Soit une image de projection inconnue décrite par le point v, qui est son descripteur de Fourier. Nous ne connaissons pas l'appartenance de l'image, ni son angle de vue. Nous calculons les distances du point v à tous les 120 (ou 180) segments de droite. Cette distance est bien entendue définie comme la longueur de la ligne normale au segment du point v. La distance la plus courte parmi les 120 (ou 180) distances donne la classe de l'image v. Pour trouver la distance minimale entre le point v et une courbe nous cherchons dans la courbe d'abord le point d'inflexion le plus rapproché au point v et nous calculons ensuite les distances entre le point v et les deux segments de droites partant de ce point d'inflexion.

Soit d_k, la distance minimale du point v à la courbe ρ_k, où $k = 1, 2,...,5$, et $d_v < d_k$ pour $v \neq k$, nous concluons que le point v décrit une image de projection du véhicule qui est décrit par la courbe ρ_v. Cette décision est prise par une étape finale du type "gagnant-prend-tout" dans l'algorithme.

Pour le cas où l'on procède à l'apprentissage à l'aide de 36 images de projection espacées entre elles de 10 degrés, les résultats de la classification pour un total de 180 images sont présentés au Table 1, selon la dimension (DIM) des descripteurs de Fourier utilisés. Seuls les nombres de fausses détections et les taux de classification réussie sont donnés. Le Table 2 donne les résultats pour un apprentissage à l'aide de 24 images de projection également espacées de 15 degrés. Mentionnons également que les résultats obtenus avec les DFs normalisés par leurs variances, tel que vu précédemment, n'apportèrent aucune amélioration.

	DIM=10	DIM=20	DIM=30	DIM=50	DIM=75	DIM=100
BTR-152	10 (86%)	7 (90%)	6 (92%)	5 (93%)	4 (94%)	4 (94%)
CHIEF	4 (94%)	3 (96%)	3 (96%)	3 (96%)	2 (97%)	1 (99%)
GRIZZLY	8 (89%)	7 (90%)	6 (92%)	6 (92%)	5 (93%)	4 (94%)
LEOPARD	1 (99%)	1 (99%)	1 (99%)	1 (99%)	1 (99%)	1 (99%)
PT-76	10 (86%)	13 (82%)	11 (85%)	11 (85%)	13 (82%)	8 (89%)
TOTAL	33 (91%)	31 (91%)	27 (93%)	26 (93%)	25 (93%)	18 (95%)

Table 1. Nombre de fausses détections pour chaque véhicule selon la dimension des descripteurs de Fourier utilisés. Entre parenthèses, le pourcentage d'efficacité de la reconnaissance est donnée. Résultats pour 10° d'incrémentation entre les rotations hors-plan.

	DIM=10	DIM=20	DIM=30	DIM=50	DIM=75	DIM=100
BTR-152	14 (81%)	15 (79%)	10 (86%)	11 (85%)	11 (85%)	9 (88%)
CHIEF	3 (96%)	5 (93%)	3 (96%)	4 (94%)	3 (96%)	1 (99%)
GRIZZLY	11 (85%)	7 (90%)	8 (89%)	7 (90%)	8 (89%)	5 (93%)
LEOPARD	3 (96%)	8 (89%)	8 (89%)	9 (88%)	8 (89%)	6 (92%)
PT-76	16 (78%)	15 (79%)	16 (78%)	15 (79%)	16 (78%)	15 (79%)
TOTAL	47 (87%)	50 (86%)	45 (88%)	46 (87%)	46 (87%)	36 (90%)

Table 2. Nombre de fausses détections pour chaque véhicule selon la dimension des descripteurs de Fourier utilisés. Entre parenthèses, le pourcentage d'efficacité de la technique de reconnaissance est donnée. Résultats pour 15° d'incrémentation entre les rotations hors-plan.

Les figures suivantes montrent les fausses détections. La coordonnée verticale donne les distances

$$d_{kv}' = d_k - d_v$$

où d_v est la distance minimale du point v à la courbe ρ_v, qui représente l'objet v, et d_k est la distance entre le point v à la courbe ρ_k, où l'objet k n'appartient pas à la classe de l'objet v. A la Table 1, on décrit les différents symboles pour les classes k associés aux points représentés dans les figures. Les distances négatives correspondent aux fausses détections dans la classification basée sur le principe de la distance minimale. L'axe des abscisses est l'angle de projection de l'image d'entrée.

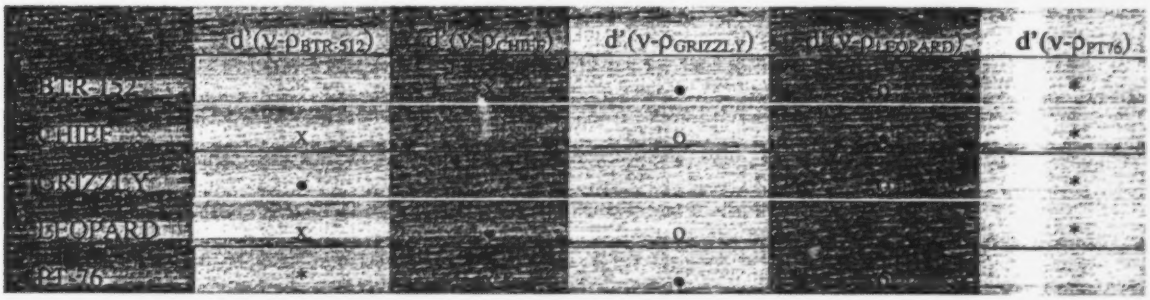


Table 3. Symboles utilisés dans les figures donnant les fausses détections dans le processus de reconnaissance.

On voit dans les figures que les fausses détections se produisent particulièrement quand les images sont des projections à 0, 180 et 360 degrés. Pour les vues de face ou de dos, les cinq véhicules à étudier ont des profils similaires. On voit aussi que le nombre de fausses détections diminue quand le nombre de coefficients d'amplitude dans les descripteurs de Fourier augmente.

Les figures suivantes montrent les fausses détections. La coordonnée verticale donne les distances

$$d'_{kv} = d_k - d_v$$

où d_v est la distance minimale du point v à la courbe ρ_v, qui représente l'objet v, et d_k est la distance entre le point v à la courbe ρ_k, où l'objet k n'appartient pas à la classe de l'objet v. A la Table 1, on décrit les différents symboles pour les classes k associés aux points représentés dans les figures. Les distances négatives correspondent aux fausses détections dans la classification basée sur le principe de la distance minimale. L'axe des abscisses est l'angle de projection de l'image d'entrée.

	$d'(v\text{-}\rho_{BTR\text{-}512})$	$d'(v\text{-}\rho_{CHIEF})$	$d'(v\text{-}\rho_{GRIZZLY})$	$d'(v\text{-}\rho_{LEOPARD})$	$d'(v\text{-}\rho_{PT76})$
BTR-152		X	•	o	*
CHIEF	X		o	•	*
GRIZZLY	•	X		o	*
LEOPARD	X	•	o		*
PT-76	*	X	•	o	

Table 3. Symboles utilisés dans les figures donnant les fausses détections dans le processus de reconnaissance.

On voit dans les figures que les fausses détections se produisent particulièrement quand les images sont des projections à 0, 180 et 360 degrés. Pour les vues de face ou de dos, les cinq véhicules à étudier ont des profils similaires. On voit aussi que le nombre de fausses détections diminue quand le nombre de coefficients d'amplitude dans les descripteurs de Fourier augmente.

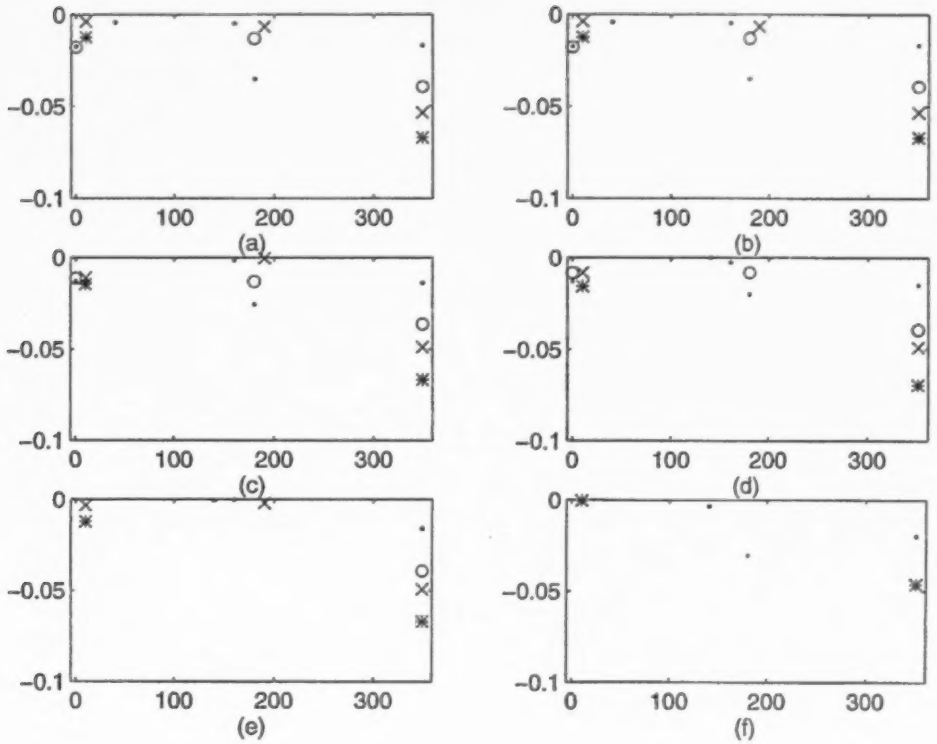


Figure 68. Graphes des distances euclidiennes représentant les fausses détections (d'<0) entre les 72 vues d'un véhicule BTR-152, données par leur angle sur l'axe des abscisses, et les courbes dans l'espace des caractéristiques pour un apprentissage à tous les 10 degrés et pour la dimension: (a) n=10 ; (b) n=20 ; (c) n=30 ; (d) n=50 ; (e) n=75 ; (f) n=100.

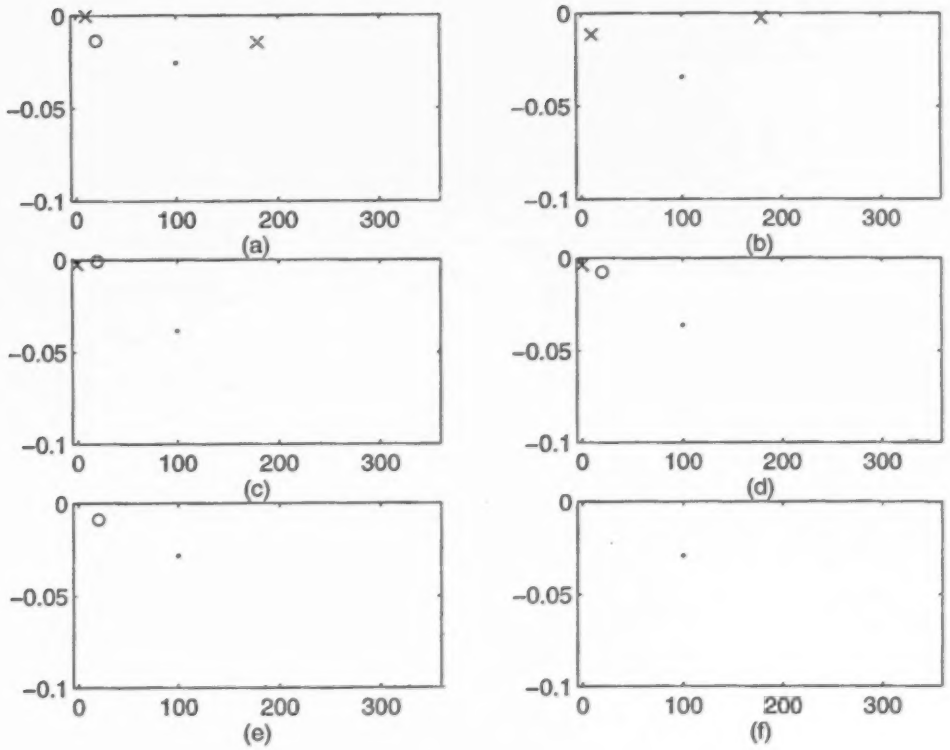


Figure 69. Graphes des distances euclidiennes représentant les fausses détections (d'<0) entre les 72 vues d'un char CHIEF et les courbes dans l'espace des caractéristiques pour un apprentissage à tous les 10 degrés et pour la dimension: (a) n=10 ; (b) n=20 ; (c) n=30 ; (d) n=50 ; (e) n=75 ; (f) n=100.

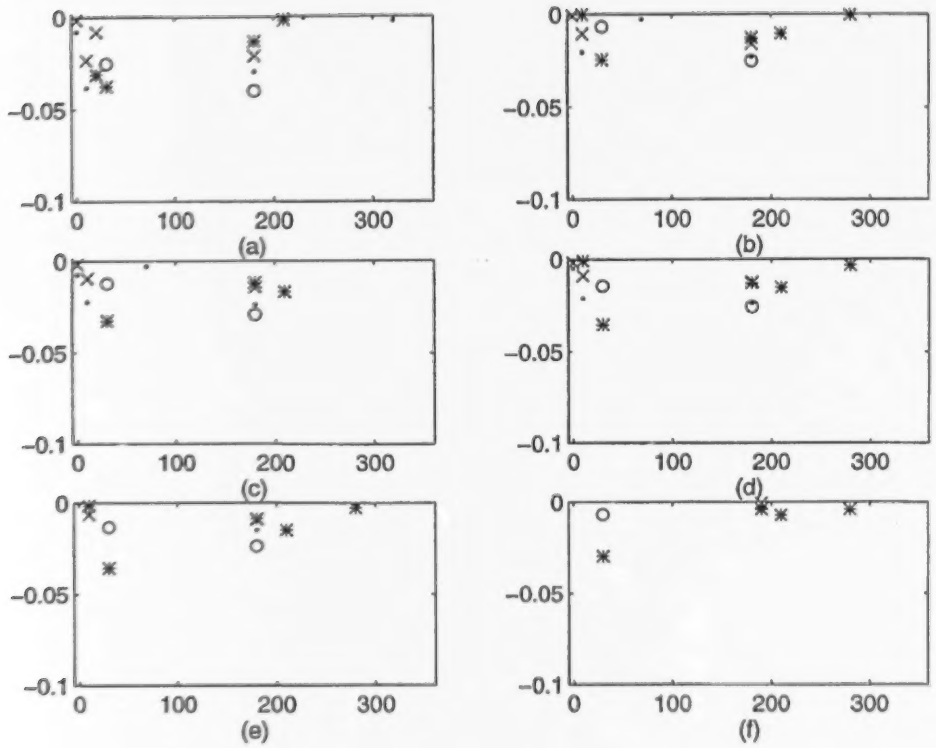


Figure 70. Graphes des distances euclidiennes représentant les fausses détections (d'<0) entre les 72 vues d'un véhicule GRIZZLY, données par leur angle sur l'axe des abscisses, et les courbes dans l'espace des caractéristiques pour un apprentissage à tous les 10 degrés et pour la dimension: (a) n=10 ; (b) n=20 ; (c) n=30 ; (d) n=50 ; (e) n=75 ; (f) n=100.

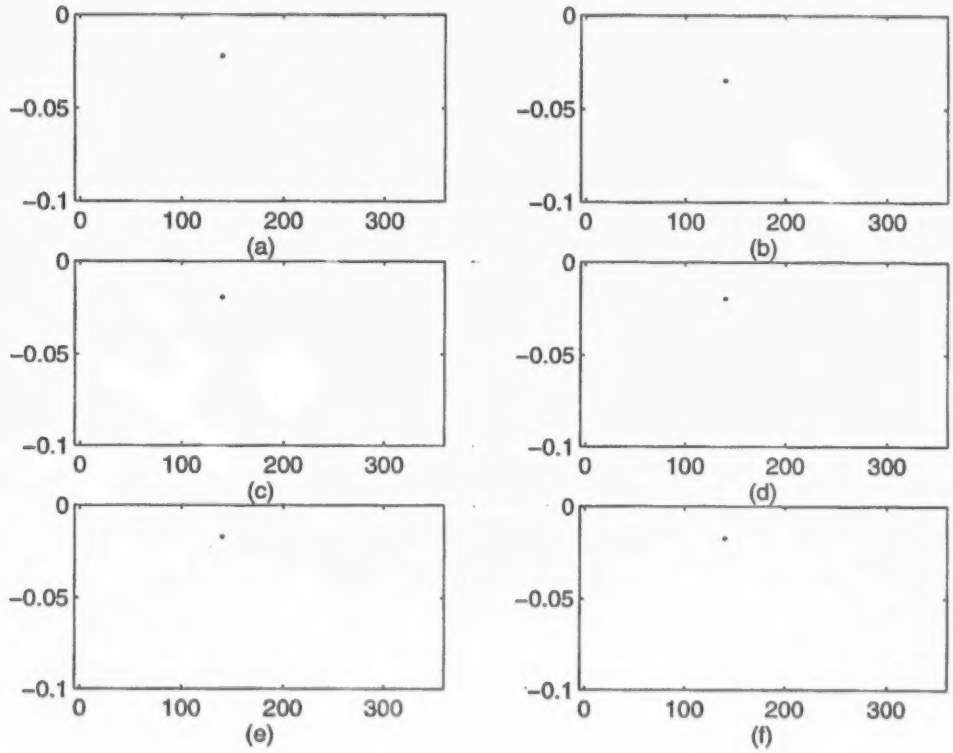


Figure 71. Graphes des distances euclidiennes représentant les fausses détections (d'<0) entre les 72 vues d'un char LEOPARD et les courbes dans l'espace des caractéristiques pour un apprentissage à tous les 10 degrés et pour la dimension: (a) n=10 ; (b) n=20 ; (c) n=30 ; (d) n=50 ; (e) n=75 ; (f) n=100.

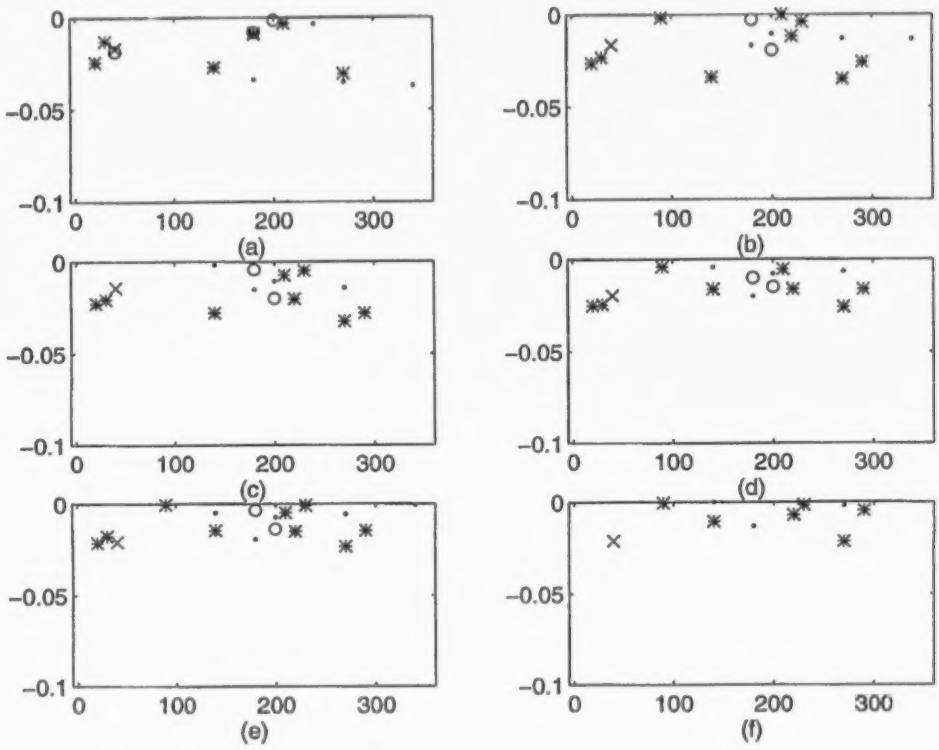


Figure 72. Graphes des distances euclidiennes représentant les fausses détections (d'<0) entre les 72 vues d'un véhicule PT-76, données par leur angle sur l'axe des abscisses, et les courbes dans l'espace des caractéristiques pour un apprentissage à tous les 10 degrés et pour la dimension: (a) n=10 ; (b) n=20 ; (c) n=30 ; (d) n=50 ; (e) n=75 ; (f) n=100.

14. Conclusion

Ce projet visait la reconnaissance automatique de cibles, problème qui diffère de la reconnaissance des formes commerciale et de la vision robotique. Entre autres, nous devions procéder à la reconnaissance d'objets 3-D à partir d'images 2-D de projection. De plus, nous avions à travailler à partir d'images prises par une caméra infrarouge. Dans l'approche proposée par Y.Sheng et S. Deschênes, nous avons travaillé sur ces deux aspects du problème à l'aide d'une procédure en quatre étapes.

Le rehaussement et la segmentation de l'image est une étape critique dans la reconnaissance automatique de cibles dans un environnement non-controlé. Nous avons procédé à l'égalisation de l'histogramme des images. Cette opération s'est avérée pertinente pour des images infrarouges de véhicules car la température et la luminance de l'image varie durant le fonctionnement du véhicule. Le filtre médian fut utilisé dans le prétraitement pour réduire le bruit thermique qui affecte la qualité du résultat lors de l'extraction du contour de l'image. Nous proposons un rehaussement des contrastes locaux dans le prétraitement, un processus pertinent pour aider à la fermeture du contour. La segmentation de l'image n'a pas été étudiée car il n'y avait pas d'objets parasites dans le fond.

Nous proposons une approche utilisant le contour de l'objet. Ce dernier contient l'information essentielle de l'objet et il est moins sensible aux changements de la distribution lumineuse, associés au fonctionnement du véhicule. Nous avons étudié différentes méthodes d'extraction du contour: les filtres de Prewitt, Sobel et Frei-Chen, le filtre optimal de Canny et la transformée en ondelette. Nous avons choisi l'ondelette de spline quadratique, qui est la première dérivée de la spline cubique, car cette ondelette est une approximation de l'un des détecteurs optimaux de contours: la première dérivée d'une gaussienne. La transformée en ondelette avec la spline quadratique est réalisée avec une paire de filtres de miroirs quadratiques dans le cadre de l'analyse multirésolutionnelle. Ceci assure la rapidité du traitement numérique utilisé. Nous avons donc une performance de détecteur de contour presqu'optimale et une réalisation numérique très rapide. L'algorithme pour la transformée avec cette ondelette a été développé dans le cadre de ce contrat. La transformée en ondelette permet de choisir la bande de fréquence dans laquelle l'image est lissée et les dérivées sont calculées. Elle permet aussi, et ce malgré le bruit, d'extraire le contour en examinant son comportement à travers les multiples résolutions. Une restitution de l'image à partir des contours à de multiples résolutions démontre la puissance de la transformée en ondelette au niveau de la conservation de l'information.

Sur le contour d'une image 2-D, l'intensité est discontinue dans une direction mais continue dans la direction perpendiculaire. Un tracé du contour à partir de l'amplitude et de la phase de la transformée en ondelette s'impose. Le détecteur de contour par la transformée en ondelette nécessite un seuillage et dépend des contrastes locaux de l'image. A ce niveau, nous avons rencontré certains problèmes pour la fermeture du contour car les contrastes étaient très faibles pour certaines parties de l'objet. Nous avons effectué des interventions non-automatisées, mais ce problème aurait pu être évité par un rehaussement du contraste lors du prétraitement.

Notre procédure produit des contours d'épaisseur d'un pixel, qui est facilement paramétrisés numériquement. Le descripteur de Fourier, basé sur les angles tangents de la courbe en fonction de la longueur d'arc, a été utilisé. La théorie nous assure que le DF est invariant sous changement

d'échelle, rotation dans le plan et translation. Nous n'avons cependant pas testé ces invariances expérimentalement.

Pour la classification, nous avons adopté le modèle de trajectoire dans l'espace des caractéristiques proposé par Casasent. Ce dernier utilise la notion de distance minimale. La base de ce modèle est l'hypothèse selon laquelle la variation de l'image de projection est graduelle, continue et lissée, quand l'objet 3-D subit une rotation hors du plan de l'image. Cette hypothèse se traduit par une interpolation linéaire dans l'espace de caractéristiques de ce modèle. Ce dernier est simple à réaliser et notre expérience montre l'efficacité du modèle.

Nos résultats expérimentaux montrent un taux de réussite de 95% et 90% pour les images d'apprentissage de projections à intervalles de 10 et 15 degrés respectivement. Ces résultats furent obtenus en testant 360 images decrivant 5 objets par leurs projections à intervalle régulier de 5 degrés, avec 100 coefficients d'amplitude du descripteur de Fourier pour chaque image. Les fausses détections se produisent surtout quand les images sont prises de face ou de dos. Les cinq véhicules étudiés ont des profils similaires sous cet angle de projection. La reconnaissance des chars d'assaut CHIEF et LEOPARD est très bien réussie avec une seule fausse détection parmi 360 images à classifier car ces véhicules donnent des profils très particuliers.

Une des difficultés principales de la reconnaissance des objets 3-D à partir de projections 2-D est le trop grand nombre d'images entrant dans l'apprentissage. Dans nos expériences et celles présentées par Casasent, seules des rotations hors-plan de l'image, sans changement en élevation, ont été considérées. Le cas complet de reconnaissance 3-D augmente encore le nombre d'images à classifiér de N (= 36 ou 24) à N^2. Les trajectoires deviennent alors des surfaces dans l'espace des caractéristiques. La performance d'un tel algorithme est inconnue.

L'approche des corrélateurs optiques demeure une solution attrayante pour l'application à la reconnaissance des objets 3-D à partir d'images de projection. Des corrélateurs sont d'ailleurs maintenant disponibles et ils réussissent à réaliser mille corrélations par seconde. Chaque filtre composite peut combiner les informations de 100 à 200 images. Cette technique permet donc de comparer l'image d'entrée avec les cent mille images stockées, sans qu'aucune segmentation ne soit nécéssaire. Les deux techniques clefs du corrélateur optique sont la conception des filtres invariants et le modulateur spatial de phase et d'amplitude complexe, développés dans nos laboratoires depuis plusieurs années.

Conclusion générale (COPL)

15. Conclusion générale (COPL)

Cette conclusion complète les conclusions qui terminent les sections sur les deux méthodes retenues. Nous avons choisi, en accord avec le responsable scientifique, deux méthodes prometteuses pour la reconnaissance de cibles ayant subi des rotations hors-plan. Avec l'accord du responsable, et parce que la base de données ne comprenait que des véhicules ayant subi des rotations horizontales, nous n'avons considéré que ce type de rotations. La base de données comprenait des chars et des transporteurs de troupes, et les résultats montrent qu'il est possible, en utilisant les méthodes étudiées ici, de les reconnaître avec des rotations arbitraires et avec une large plage d'échelles avec un degré de certitude pouvant atteindre 100% pour les images fournies.

Nous considérons que, malgré les montants relativement faibles accordés pour ce sous-contrat, nous avons obtenu beaucoup de résultats. En rétrospective, on peut cependant identifier quelques éléments qui auraient permis d'effectuer cette recherche avec plus de rigueur et possiblement de pousser plus loin. La base de données comprend des images infrarouges video de pauvre qualité. Bien que ces images donnent une idée plus réaliste de ce qu'on pourrait trouver dans les conditions réelles, étant donné la complexité du problème et le peu de résultats obtenus avant nous par d'autres chercheurs, il aurait été plus "scientifique" de commencer par des images de meilleure qualité. En effet, quand nous avons commencé à avoir des difficultés, nous avons perdu beaucoup de temps pour déterminer si les difficultés étaient dues aux limitations de la méthode, à la pauvre qualité des images, ou à des artifices introduits par les méthodes numériques. En commençant avec des images sans bruit, on aurait pu déterminer avec précision les limites de la méthode elle-même, pour ensuite déterminer à l'aides des images réalistes à quel point les erreurs sont augmentées à cause de la qualité des images. Il aurait été avantageux d'avoir une base de donnée avec non seulement des orientations différentes, mais aussi avec des échelles différentes. En effet, nous avons dû changer les échelles des véhicules au moyen d'algorithmes numériques, et nous avons constaté que différentes méthodes donnaient des différences significatives pour la reconnaissance. Nous pensons (mais n'avons aucun moyen de le vérifier) qu'une proportion des erreurs provient du repliement spectral causé par les algorithmes de changement d'échelle, et que celui-ci diffère du repliement spectral qui serait causé par des images ayant été photographiées à des échelles différentes. Il est donc difficile de prédire avec certitude la performance réelle dans des conditions réelles des méthodes étudiées. Finalement, notons que l'une des deux méthodes étudiées se prête à une réalisation optique. Pour cette fin, on pourrait utiliser des composantes peu dispendieuses développées ici avec de l'optique binaire, et qui pourraient être améliorées en utilisant des techniques d'optique diffractive. Les directions futures de cette recherche s'imposent donc d'elles-même on a vu un progrès considérable au cours des deux dernières années, et les résultats sont très prometteurs, mais il reste encore beaucoup de questions qui méritent d'être étudiées.

[1] S. J. Perantonis et P. J. G. Lisboa, "Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifiers", IEEE trans. on neural net., Vol.3 No.2, 241-251 (1992).

[2] M. Fukumi, S. Omatu, F. Takeda et T. Kosaka, "Rotation-invariant neural pattern recognition system with application to coin recognition", IEEE trans. on neural net., Vol.3, No.2, 272-279 (1992).

[3] L. Spirkovska et M. B. Reid, "Robust position, scale, and rotation invariant object recognition using higher-order neural networks", Pattern Recognition, Vol.25, No.9, 975-985 (1992).

[4] R. J. Prokop et A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition", Graphical models and image Processing, Vol.54, No.5, 438-460 (1992).

[5] N. T. Nguyen et D. X. Xu, "Scale based algorithm for recognition of blurred planar objects", IEE Proceedings, Vol.135, No.6, 307-311 (1988).

[6] S. Z. Li, "Matching: invariant to translations, rotations and scale changes", Pattern recognition, Vol.25, No.6, 583-594, (1992).

[7] A. Bryant et J. Bryant, "Recognizing shapes in planar binary images", Pattern Recognition, Vol.22, No.2, 155-164 (1989).

[8] K. Cho et S. M. Dunn, "Hierarchical local symmetry: 2-D shape representation", SPIE, Vol. 1192 Intelligent Robots and computer vision VII: algorytyhms and techniques, 208-217, (1989).

[9] J. Xu et Y. Yang, "Generalized multidimensional orthogonal polynomials with applications to shape analysis", IEEE trans. on pattern anal. and machine intelligence, Vol.12, No.9, 906-913, (1990).

[10] Y. Sheng et C. Lejeune, "Invariant pattern recognition using Fourier-Mellin transforms and neural networks", J. Optics, Vol.22, No.5, 223-228, (1991).

[11] F. T. S. Yu, Y. Li et T. Lu, "Application of moment invariant pattern recognition to optical neural net", Optik, 89, No.2, 55-58 (1991).

[12] S. Chang, H. H. Arsenault et D. Liu, "Invariant optical-pattern recognition based on contour bank", Applied optics, Vol.33, No.14, 3976-3085, (1994).

[13] E. Elizur et A. A. Friesem, "Rotation-invariant correlation with incoherent light", Applied optics, Vol.30, No.29, 4175-4178, (1991).

[14] J. Garcia, C. Ferreira, J. Campos, S. Bosch, M. S. Millan et M. J. Yzuel, "Multichannel rotation-invariant pattern recognition for polychromatic objects using circular harmonic filters", Optics communications, 91, 425-432, (1992).

[15] M. B. Reid, P. W. Ma, J. D. Downie et E. Ochoa, "Experimental verification of modified synthetic discriminant function filters for rotation invariance", Applied optics, Vol.29, No.8, 1209-1214, (1990).

[16] G. Ravichandran et D. P. Casasent, "Genralized in-plane rotation-invariant minimum average correlation energie filter", Optical engineering, Vol.30, No.10, 1601-1607, (1991).

[17] Burns J. B., Weiss R. S., Edward M. R., "View variation of point-set and line-segment features", Proc. DARPA Image Undestanding Workshop, pp. 650-659, Pittsburg, April 1990.

[18] Weinshall D., "Model-based invariants for 3-D vision", International Journal of Computier Vision, pp. 27-42, 1993.

[19] Forsyth D., Mundy J. L., Zisserman A., Coelho C., Heller A., Rothwell C., "Invariant descriptors for 3-D object recognition and pose", IEEE Trans. Patt. Anal. Mach. Intell., 13: 971-991, 1991.

[20] Ullman S., Basri R., "Recognition by linear combinations of models", IEEE Trans. Patt. Anal. Mach. Intell., 13: 992-1006, 1991.

[21] C. E. Daniell et.al., "Artificial neural networks for automatic target recognition", Opt. Eng. **31**, 2521-2531 (1992).

[22] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biol. Cybernetics, **36**, 193-202 (1980).

[23] J. G. Verly et.al., "Model-based system for automatic target recognition from forward-looking laser-radar", Opt. Eng. **31**, 2540-2552 (1992).

[24] J. E. Hummel et I. Biederman "Dynamic binding in a neural network for shape recognition", Psychol. Rev. **99**, 480-517 (1992)

[25] W. A. Thoet et.al. "ANVIL neural network program for three-dimensional automatic target recognition", Opt. Eng. **32**, 2532-2539 (1992).

[26] Johnson J. L. "Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images", Applied Optics, Vol 33, No 26, pp. 6239-6253, 1994.

[27] L. Neiberg et D. P. Casasent, "Feature Space Trajectory (FST) Classifier Neural Network", Proceeding SPIE, Vol 2353, 276-292 (1994)

[28] G. G. Lendaris et G. L. Stanley, "Diffraction-pattern sampling for automatic pattern recognition", Proc. of the IEEE, Vol.58, No.2, 198-216, (1970).

[29] N. George, S. Wang et D. L. Venable, "Pattern recognition using the ring-wedge detector and neural-network software", Optical pattern recognition II, SPIE Vol.1134, (1989).

[30] W. K. Pratt, "Digital image processing", Chap.10, 2nd. ed. (John Wiley, Ner York, 1991).

[31] S. Grossberg et E. Mingolla, "Neural dunamics of surface perception: boundary webs, illuminants, and shape from shading", Comput. Vision, Graph. Image Process., **37**, 116-165 (1987)

[32] J. Canny, "A computational approach to edge detection", IEEE Trans. on Patt. anal. and Machine Intell., **PAMI-8**, 679-698 (1986).

[33] Y. Sheng, "Wavelet transform", in "The Transforms and Applications Handbook", A.D. Poulariskas Ed. Chap. 10, (CRC Press, Boca Raton, 1995)

[34] D. Marr et E. Hildreth, "Theory of edge detection", Proc. R. Soc. Lond., **B 207**, 187-217 (1980).

[35] Y. Sheng, D. Roberge et H. Szu, "Optical wavelet transform", Opt. Eng., **31**, 1840-1845 (1992).

[36] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Trans. Pattern Anl. Machine Intell. PAMI-31, 674-693 (1989).

[37] S. Mallat, W. L. Hwang, "Singularity detection and processing with wavelet", IEEE Tran. on Inf. Theory, **38**, 617-641 (1992).

[38] F. Sadjadi, "Object recognition using coding schemes", Opt. Eng. **32**, 2580-2583 (1992).

[39] Y. Sheng, H. H. Arsenault, "Experiments on the pattern recognition using invariant Fourier-Mellin descriptors", J. Opt. Soc. Am. A, **3**, 6, 771-776 (1986).

[40] M. K. Hu, "Visual pattern recognition by moment invariants" IRE Trans. Inf. Theory, **IT-8** 179-187 (1962).

[41] Y. Sheng and L. Shen, "Orthogonal Fourier-Mellin moments for invariant pattern recognition", J. Opt. Soc. Am. A, **6**, 1748-1757 (1994).

[42] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane curves" IEEE Trans. on Computers, Vol. c-**21**, 3, 269-277 (1972).

[43] W. J. Shomar et T. Y. Young, "Three-dimensional shape recovery from line drawings", Chap.3 in the book "Handbook of pattern recognition and image processing: computer vision" Vol.2, T. Y. Young Ed. (Academic, San Diego, 1994).

[44] K. Arbter, W. E. Snyder, H. Burkhardt et G. Hirzinger, "Application of Affine-Invariant Fourier Descriptors to Recognition of 3-D Objects", IEEE Trans. on Patt. anal. and Machine Intell., **PAMI**, Vol. 12, No 7, 640-647,July 1990

[45] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biol. Cybernetics, **36**, 193-202 (1980).

[46] D. E. Rumelhart, G. E. Hinton et R. J. Williams, "Learning representations by back-propagating errors", Nature, **323**, 533-536 (1986).

[47] T. Poggio et S. Edelman, "A network that learns to recognize three-dimensional objects", Nature, **343**, 263-266 (1990).

[48] H. C. Longuet-Higgins, Nature **292**, 133-135 (1981).

Conclusion Générale

Conclusion Générale

Les expérimentations numériques et optiques qui ont été menées au cours du projet permettent de tirer des conclusions d'ordres général et particulier. Les conclusions générales seront tout d'abord présentées, suivies du détail des conclusions particulières à chacune des parties du projet. On trouvera de plus à la fin de cette conclusion les principales perspectives qui découlent des travaux effectués.

La première conclusion générale que l'on retient du projet est que la reconnaissance d'objets 3D à partir de la seule information provenant d'une caméra est un problème certe complexe, mais soluble du moins dans un contexte simple et un environnement structuré. C'est pour cette raison que les travaux furent concentrés autour d'images tirées d'une série limitée d'objets en rotation autour d'un axe vertical.

En second lieu, il est ressorti que le prétraitement des images (rehaussement du contraste, suppression des bruits et segmentation de l'image) joue un rôle capital dont dépend grandement la performance de toutes les méthodes de reconnaissance automatique qui ont été considérées. Une étude plus approfondie des problèmes liés au prétraitement des images en RAC serait certainement bénéfique pour l'amélioration de la performance, mais surtout de la robustesse des traitements subséquents. Bien que quelques méthodes de prétraitement aient été utilisées en simulation numérique, il a été jugé préférable dans un premier temps de caractériser le modèle d'essai optique à l'aide d'images synthétiques non bruitées, telles que nous les auraient fournies un module de prétraitement idéal.

Au chapitre des conclusions particulières, nous sépareront le projet en deux parties: l'évaluation des méthodes par simulation numériques et la réalisation du modèle d'essai. Les différentes méthodes mises à l'essai par les équipes de l'INO et de l'Université Laval seront passées en revue, il s'agit des méthodes des sections en pointe, des descripteurs de Fourier, et des signatures CHC/CHCA. Elles ont toutes fait l'objet de simulations numériques, la dernière a été retenue pour la réalisation du modèle d'essai optique.

- Evaluation des méthodes par simulation numérique:

Les simulations numériques effectuées ont montré que les signatures CHC et CHCA sont invariantes (fluctuations inférieures à 3%) sous rotation dans le plan et sous échelle, et cela malgré la pixellisation des images et leur faible résolution. La mise en oeuvre numérique de la méthode CHC permet d'obtenir des signatures comportant peu de variabilité sous les rotations dans le plan et changement d'échelle. Combinée à des réseaux de neurones, elle a permis la reconnaissance 3D des objets de test.

Cependant, les simulations ont aussi démontré que cette méthode est sensible à la qualité du prétraitement, et qu'afin d'obtenir l'invariance, il faut que l'objet soit extrait de l'arrière scène. Cette opération doit être effectuée sans trop détériorer les contours de la silhouette de l'objet. La méthode des CHC est aussi sensible aux erreurs de centrage. Dans le contexte contrôlé de nos expérimentations, le recentrage de l'objet par l'entremise du calcul du centre géométrique a fourni une stabilité suffisante à l'obtention de bons résultats. Toutefois dans un contexte plus réaliste l'impact d'erreurs sur la détermination du centre, dus par exemple à l'illumination ambiante ou aux occlusions, n'a pas fait l'objet d'une évaluation systématique. Il pourrait y avoir là un intérêt à retourner à l'approche proposée initialement, à savoir procéder à l'extraction des CHC sur la version spectrale (TF) de l'objet plutôt que sur l'objet lui-même.

La méthode des descripteurs de Fourier sur les contours est la plus complexe de celles étudiées dans le cadre de ce projet. C'est une méthode en principe invariante sous échelle, translation et rotation dans le plan mais elle ne se prête pas bien à une mise en oeuvre optique, en raison de la nature des calculs impliqués. L'une des faiblesses de la méthode réside dans la nécessité d'effectuer l'extraction précise du contour de la silhouette de l'objet. Dans un contexte de RAC, l'extraction de la silhouette est une opération difficile et délicate. Les résultats seront d'autant meilleurs que l'extraction de la silhouette sera fiable et précise. L'étude de cette méthode a donné lieu à des travaux intéressants au sujet de l'utilisation de la transformation en ondelette pour l'extraction des contours.

Enfin, la méthode des sections en pointe a fourni des résultats intéressants. Elle est toutefois très exigeante en ce qui a trait à la capacité de calcul puisqu'elle nécessite le calcul

du module carré de la TF de l'objet. C'est pour cette raison que la mise en oeuvre optique de la méthode serait avantageuse. Pour ce faire, il faudrait disposer d'un système optique de très bonne qualité, qui permette le calcul du module de la TF sans changement trop appréciable de l'aspect avec les translations et rotations dans le plan. Un tel système optique nécessiterait également le création de détecteurs à configuration particulière (gros pixels en forme de pointe). Les simulations ont montré que la méthode se comporte assez bien pour la reconnaissance d'objet 3D dans un contexte simple pour lequel les rotations sont limitées autour d'un seul axe (vertical). La méthode nécessite, comme les autres, des opérations de prétraitement efficaces. La méthode proposée est en principe invariante sous changement d'échelle mais pas sous rotation dans le plan. L'évaluation des modifications à apporter afin de la rendre invariante sous les rotations dans le plan n'a pas été effectuée dans le cadre de ce projet.

- Réalisation du modèle d'essai:

En ce qui a trait à la réalisation du modèle d'essai, l'extraction optique des CHC (et des CHCA) implique des mesures de précision sur des valeurs des rosettes en 1+cos et en 1+sin. Deux approches différentes ont été utilisées: un système à multiplexage spatial des rosettes en éclairage cohérent et un système à multiplexage temporel des rosettes en éclairage incohérent.

Les résultats obtenus avec le système incohérent à deux ACL sont suffisamment bons pour fournir des signatures CHC et CHCA invariantes, qui permettent l'entraînement d'un réseau de neurones à reconnaître des objets indépendamment de leur orientation dans le plan (rotation dans le plan) et de leur échelle.

Quant au système cohérent, basé sur un réseau de diffraction, il n'a pas donné de résultats finaux en raison de problèmes rencontrés dans la fabrication des masques. Les expérimentations ont toutefois permis de faire ressortir les inconvénients d'un tel système cohérent pour l'extraction des caractéristiques morphologiques. En effet, les architectures optiques nécessitant un éclairement cohérent impliquent des difficultés importantes de mise en oeuvre en raison du bruit. Le bruit dans un système optique cohérent est causé principalement par la poussière qui se dépose sur les surfaces optiques. De plus pour éviter

l'apparition de franges d'interférences et d'autres effets indésirables, les systèmes cohérents nécessitent des composantes optiques de très bonne qualité. L'utilisation de l'optique cohérente devrait se réduire uniquement au calcul de transformée de Fourier (TF) et des opérations pouvant se ramener à des TF. Pour les autres types de calculs il semble préférable d'utiliser les technologies de l'optique incohérente de concert avec le numérique.

- Perspectives:

Il se dégage de l'ensemble des travaux effectués dans ce projet plusieurs perspectives intéressantes, dont voici quelques unes qui représentent selon nous les avenues de recherches préférentielles dans une suite éventuelle des travaux.

En ce qui a trait au développement futur du modèle d'essai que nous avons réalisé pour l'extraction optique des CHC et des CHCA, nous sommes d'ores et déjà en mesure de proposer une évolution prometteuse et intégrée (présentée à la fin de la section 3.3.2). Celle-ci présente non seulement l'avantage d'être compacte, mais également la capacité de réaliser l'extraction de quatre valeurs 1+cos et/ou 1+sin à l'aide d'un seul ACL de format VGA. Il deviendrait alors possible de construire un processeur optique parallèle à peu de frais, en combinant quelques-uns de ces dispositifs.

Encore au plan des réalisations optiques, nous avons mentionné qu'il nous apparait utile de revenir à des architectures basées sur le module carré de la TF de l'objet, plutôt que sur l'image de l'objet lui-même. Ceci implique la réalisation d'un système optique de haute qualité permettant le calcul du module carré de la TF qui ne change pas d'aspect lorsqu'il y a translation ou rotation dans le plan. Le développement d'un tel système optique demanderait également que l'on se penche sur l'étude de moyens pour modifier la TF à l'aide d'un masque pour éviter la saturation de la caméra CCD, ou encore d'étudier la possibilité d'utiliser de nouveaux types de caméra à plage de sensibilité étendue (ex. caméra CMOS).

Ayant maintenant une vision plus précise des principales méthodes applicables au problème de la RAC, ainsi que de leur performances en simulation numérique, il serait intéressant d'étudier la possibilité de fusion de certaines de ces méthodes. La combinaison

pour obtenir de meilleurs résultats parait prometteuse puisque les méthodes étudiées impliquent la classification sur la base d'information synthétisée. Par exemple dans le cas de la méthode des CHC, la classification des objets doit se faire à partir de quelques coefficients CHC seulement. L'enrichissement de l'information fournis par plusieurs méthodes permet souvent des performances supérieures à ce que chacune des méthodes fournit individuellement. Par exemple, il serait intéressant de combiner la méthode des CHC/CHCA avec celle des sections en pointe.

Enfin, il ne faudrait pas omettre de mentionner dans les perspectives ce qui touche à l'approche par corrélation optique, qui demeure un domaine d'intérêt pour la RAC. Celle-ci n'a été étudiée que sommairement dans la première partie de ce projet, en raison surtout des limitations de l'état de l'art à ce moment. C'est une approche qui nécessite aussi bien une réflexion algorithmique (étant donné le nombre important de filtres impliqués), que le développement et l'intégration de composants optoélectroniques haut-de-gamme au sein d'un corrélateur optique. La vitesse de traitement d'un corrélateur optique est principalement limitée par la cadence de rafraîchissement de ses afficheurs. Or, l'avenir des afficheurs est rempli de promesses puisque l'on assiste à des développements rapides et importants dans le domaine des afficheurs à cristaux liquides, notamment en raison du développement de la technologie basée sur les cristaux liquides ferroélectriques. D'autres approches technologiques voient actuellement le jour dans le domaine de l'affichage, telles les matrices de micro-miroirs déformables qui ont le potentiel de réaliser des cadences supérieures à ce qui est actuellement disponible.